PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/12, C07K 14/47, 16/18, G01N 33/50

A1

(11) International Publication Number: **WO 99/49042**

(43) International Publication Date: 30 September 1999 (30.09.99)

(21) International Application Number: PCT/US99/06714

(22) International Filing Date: 29 March 1999 (29.03.99)

(30) Priority Data:
09/049,792 27 March 1998 (27.03.98) US

(71) Applicant: ICOS CORPORATION [US/US]; 22021 20th Avenue, S.E., Bothell, WA 98021 (US).

(72) Inventor: HERENDEEN, Daniel, R.; 10934 N.E. 117th Place, Kirkland, WA 98034 (US).

(74) Agent: UHL, Jill, E.; Marshall, O'Toole, Gerstein, Murray & Borun, Suite 6300, 233 South Wacker Drive, Chicago, IL 60606 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.
With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.

(54) Title: HUMAN Rad1 CELL CYCLE CHECKPOINT

(57) Abstract

Disclosed are novel checkpoint polynucleotides encoding polypeptides designated Rad1, polypeptides encoded by the polynucleotides, expression constructs comprising the polynucleotides, host cells transformed or transfected with the polynucleotides, methods for producing the polypeptides, antibodies immunospecific for the polypeptides, methods to identify binding partners of the polypeptides, and methods to screen for modulators of Rad1 biological activity.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

Human Rad1 Cell Cycle Checkpoint

CROSS-REFERENCE TO RELATED APPLICATIONS

This application is a continuation-in-part of co-pending U.S. Serial Number 08/049,792, filed March 27, 1998 and from which this application claims priority.

BACKGROUND

The process of eukaryotic cell growth and division is the somatic, or mitotic, cell cycle which consists of four phases, the G1 phase, the S phase, the G2 phase, and the M phase. During the G1 (gap) phase, biosynthetic activities of the cell are carried out at a high rate. The S (synthesis) phase begins when DNA synthesis starts and ends when the DNA content of the nucleus of the cell has been replicated and two identical sets of chromosomes are formed. The cell then enters the G2 (gap) phase which continues until mitosis starts. In mitosis, the chromosomes pair and separate and two new nuclei form, and in cytokinesis, the cell splits into two daughter cells, each receiving one nucleus containing one of the two sets of chromosomes. Mitosis and cytokinesis together make up the M (mitosis) phase of the cell cycle. Cytokinesis terminates the M phase. The sequence in which the events of the cell cycle proceed is tightly regulated such that the initiation of one cell cycle event is dependent on the completion of the prior cell cycle event. This control allows fidelity in the duplication and segregation of genetic material from one generation to the next.

The term "cell cycle checkpoint" refers to the proteins, signals, processes, and feedback control that integrate discontinuous events during cellular replication in order to maintain essential dependencies within the cell cycle. See generally, Elledge, *Science 274*:1664-1672 (1996) and Nurse, *Cell 91.*:865-867 (1997). Failure of the cell cycle checkpoints predisposes individuals to, or directly causes, many disease states such as cancer, axatia telangiectasaia, embryo abnormalities, and various immunological defects associated with aberrant B and T cell development. The latter are associated with pathological states such as lupus, arthritis, and autoimmune diseases. Intense research efforts have therefore focused on identifying cell cycle checkpoints and the proteins essential for the function of the checkpoints.

Genetic analysis in the yeasts *S. pombe* and *S. cerevisiae* has identified a

number of genes important for cell cycle arrest and DNA repair essential to regulating cell cycle progression. For a review, see Carr and Hoekstra, *Trends in Cell Biology 5*:32-40 (1995). One such gene identified in both yeast species is required for a DNA damage checkpoint which arrests the cell cycle in response to DNA strand interruptions. The gene has been designated rad1+ in *S. pombe* and RAD17 in *S. cerevisiae*, referred to hereafter as Rad1. See generally, Grossenbacker-Grunder, *et al., Mut. Res. 81:*37-48 (1981); Sunnerhagen, *et al., Mol. Cell. Biol. 10:* 3750-3760 (1990); Rowley, *et al., EMBO J. 11:*1335-1342 (1992); Long, *et al., Gene 148:*155-159 (1994); Siede, *et al., Nucleic Acids Res. 24:*1669-1675 (1996); and Lydall, *et al., Science 270:*1488-1491 (1995).

Cells having mutations in Rad1 fail to either sense or appropriately respond to DNA damage and subsequently lose viability more rapidly than wild type cells after exposure to clastogenic agents or events (*e.g.*, radiation, DNA damaging agents, DNA replication inhibitors, and the like). Sensitivity to radiation can cause defects in checkpoint responses or defects in direct DNA repair reactions. The product of the yeast rad1+ gene is an approximately 37 kDa protein that has approximately 24% identity to the REC1 DNA exonuclease of *U. maydis* and to Rad17 of *S. cerevisiae*. REC1 controls genetic recombination in *U.maydis* and the *U. maydis* rec1 mutant, like *S. pombe* rad1 and *S. cerevisiae* rad17 mutants, is sensitive to radiation and hydroxyurea treatment. See generally, Holliday, *Mut. Res. 2:*557-559 (1965); Holliday, *Mut. Res. 4:*275-288 (1967); and Onel, *et al., Genetics 43:*165-174 (1996). Rad1 may have a role in DNA damage processing as it may encode a DNA binding activity, and associated exonuclease activity, participating in checkpoint surveillance and DNA processing.

Kuerbitz, *et al., Proc. Natl. Acad. Sci.(USA) 89*:74932-7495 (1992) disclosed that the tumor suppressor protein p53 is required for a G1 checkpoint and cell cycle arrest observed following DNA damage. Irradiation of cells was shown to result in increased levels of p53 leading to the transcriptional activation of p53 responsive genes. One such p53-induced target is the product of the WAF1 gene (also called p21, CIP1, and sid1). WAF1 is a member of an expanding class of cell cycle regulatory proteins termed cyclin-dependent kinase inhibitory proteins which control transition through the cell cycle. Transcriptional activation of WAF1 thus provides a

direct link between DNA damage-dependent induction of p53 and the inhibition of kinases essential for cell cycle progression. See Elledge and Harper, *Current Opinions in Cell Biology 6*:847-852 (1994). Mutations in the p53 gene are one of the most common genetic alteration in human cancers. For example, Baker, *et al., Cell 61*:759-767 (1990) reported that breast, lung, bladder, and brain tumors have been associated with the loss of chromosome 17p, the region to which the p53 gene has been localized.

At present, there is relatively little known about the molecular components of the G2 checkpoint in mammals. Caffeine is known to abrogate G2 checkpoint control [Russe, *et al., Cancer Res. 55*:1639-1642 (1995)]. Powell, *et al., Cancer Res 55*:1643-1649 (1995) also reported that analysis of cell lines, which differ only by the presence or absence of a functional p53, demonstrated preferential caffeine-enhanced sensitization to radiation in those cells lacking the p53-dependent checkpoint. Thus, potentially lethal damage is greater in cells lacking the G1 and G2 checkpoints in comparison to cells containing an intact checkpoint.

While certain cells undergo DNA damage-dependent cell cycle arrest, other cells appear to respond to DNA damage by initiating an intrinsic suicide program termed apoptosis or programmed cell death. The factors determining which process occurs are not full understood. Recent work has demonstrated an important role for p53 both in the regulation of G1 cell cycle transitions and apoptosis.

High doses of radiation and chemotherapy are used to treat tumor cells in order to damage DNA so severely that the cells will die. However, even though tumor cells having a mutation in the p53 gene are defective in a G1 checkpoint, they can still repair DNA damage induced by irradiation or chemotherapy.

Thus there exists a need in the art for identification of the mammalian proteins that are involved in the cell cycle checkpoints in order to develop therapies for the human disease states associated with defective cell cycle checkpoints and for the use of the isolated genes encoding those proteins which in themselves may be useful as therapies. Identification of these proteins, and their underlying genes, would enable the development of therapeutically useful modulators of the proteins encoded by the genes. Inhibition of a Rad1-dependent checkpoint in tumors cells could lead to therapies wherein tumor cells are incapable of repairing DNA damage, therefore sensitizing the tumor cells to DNA damaging agents. Normal cells, containing intact

G1 and G2 checkpoints, would still be able to repair DNA damage in the presence of the G2 checkpoint-specific inhibitor. Treatment of tumors with a Rad1 checkpoint-specific inhibitor followed by, or along with, radiation or chemotherapy would increase the efficacy of cell killing and thereby decrease the required doses of toxic DNA damaging agent.

SUMMARY OF THE INVENTION

In brief, the present invention provides novel Rad1 polypeptides and underlying polynucleotides. The invention includes both naturally occurring and non-naturally occurring Rad1 polynucleotides and polypeptide products thereof. Naturally occurring Rad1 products include distinct gene and polypeptide species within the Rad1 family; these species include those which are expressed within cells of the same animal as well as corresponding species homologs expressed in cells of other animals. Within each Rad1 species, the invention further provides splice variants encoded by the same polynucleotide but which arise from distinct mRNA transcripts. Non-naturally occurring Rad1 products include variants of the naturally occurring products such as analogs (*i.e.*, wherein one or more amino acids are added, substituted, or deleted) and those Rad1 products which include covalent modifications (*i.e.*, fusion proteins, glycosylation variants, Met^{-1}-Rad1, Met^{-2}-Lys^{-1}-Rad1s, Gly^{-1}-Rad1s and the like). In a preferred embodiment, the invention provides a polynucleotide comprising the sequence set forth in SEQ ID NO: 1. The invention also embraces polynucleotides encoding the amino acid sequence set out in SEQ ID NO: 2. A presently preferred polypeptide of the invention comprises the amino acid sequence set out in SEQ ID NO: 2.

The present invention provides novel purified and isolated polynucleotides (*e.g.*, DNA sequences and RNA transcripts, both sense and complementary antisense strands, including splice variants thereof) encoding the human Rad1s. DNA sequences of the invention include genomic and cDNA sequences as well as wholly or partially chemically synthesized DNA sequences. "Synthesized," as used herein and is understood in the art, refers to purely chemical, as opposed to enzymatic, methods for producing polynucleotides. "Wholly" synthesized DNA sequences are therefore produced entirely by chemical means, and "partially" synthesized DNAs embrace those wherein only portions of the resulting DNA were produced by chemical means. A preferred DNA sequence encoding a human Rad1 polypeptide is set out in SEQ ID NO: 1. The worker

of skill in the art will readily appreciate that the preferred DNA of the invention comprises a double stranded molecule, for example the molecule having the sequence set forth in SEQ ID NO: 1 along with the complementary molecule (the "non-coding strand" or "complement") having a sequence deducible from the sequence of SEQ ID NO: 1 according to Watson-Crick base paring rules for DNA. Also preferred are polynucleotides encoding the Rad1 polypeptide of SEQ ID NO: 2. The invention further embraces species, preferably mammalian and in particular mouse, homologs of the human Rad1 DNA. A preferred mouse polynucleotide of the invention is set out in SEQ ID NO: 3 which encodes the amino acid sequence set out in SEQ ID NO: 4.

The invention also embraces DNA sequences encoding Rad1 species which hybridize under highly stringent conditions to the non-coding strand, or complement, of the polynucleotide in SEQ ID NO: 1. DNA sequences encoding Rad1 polypeptides which would hybridize thereto but for the redundancy of the genetic code are contemplated by the invention. Exemplary high stringency hybridization conditions are as follows: hybridization at 42°C in 5X SSPE and 50% formamide, and washing at 50°C in 0.1X SSC. It is understood in the art that conditions of equivalent stringency can be achieved through variation of temperature and buffer, or salt concentration as described in Ausubel, *et al.* (Eds.), Protocols in Molecular Biology, John Wiley & Sons (1994), pp. 6.0.3 to 6.4.10. Modifications in hybridization conditions can be empirically determined or precisely calculated based on the length and the percentage of guanosine/cytosine (GC) base pairing of the probe. The hybridization conditions can be calculated as described in Sambrook, *et al.*, (Eds.), Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory Press: Cold Spring Harbor, New York (1989), pp. 9.47 to 9.51.

Autonomously replicating recombinant expression constructions such as plasmid and viral DNA vectors incorporating Rad1 sequences are also provided. Expression constructs wherein Rad1-encoding polynucleotides are operatively linked to an endogenous or exogenous expression control DNA sequence and a transcription terminator are also provided.

According to another aspect of the invention, host cells are provided, including prokaryotic and eukaryotic cells, either stably or transiently transformed with DNA sequences of the invention in a manner which permits expression of Rad1 polypeptides of the invention. Host cells of the invention are a valuable source of

immunogen for development of antibodies specifically immunoreactive with Rad1. Host cells of the invention are also conspicuously useful in methods for large scale production of Rad1 polypeptides wherein the cells are grown in a suitable culture medium and the desired polypeptide products are isolated from the cells or from the medium in which the cells are grown by, for example, immunoaffinity purification.

Knowledge of Rad1 DNA sequences allows for modification of cells to permit, or increase, expression of endogenous Rad1. Cells can be modified (*e.g.*, by homologous recombination) to provide increased Rad1 expression by replacing, in whole or in part, the naturally occurring Rad1 promoter with all or part of a heterologous promoter so that the cells express Rad1 at higher levels. The heterologous promoter is inserted in such a manner that it is operatively linked to Rad1-encoding sequences. See, for example, PCT International Publication No. WO 94/12650, PCT International Publication No. WO 92/20808, and PCT International Publication No. WO 91/09955. The invention also contemplates that, in addition to heterologous promoter DNA, amplifiable marker DNA (*e.g.*, ada, dhfr, and the multifunctional CAD gene which encodes carbamyl phosphate synthase, aspartate transcarbamylase, and dihydroorotase) and/or intron DNA may be inserted along with the heterologous promoter DNA. If linked to the Rad1 coding sequence, amplification of the marker DNA by standard selection methods results in co-amplification of the Rad1 coding sequences in the cells.

The DNA sequence information provided by the present invention also makes possible the development through, *e.g.* homologous recombination or "knock-out" strategies [Capecchi, *Science 244*:1288-1292 (1989)], of animals that fail to express functional Rad1, that express a variant of Rad1, or that have altered regulation of Rad1. Such animals are useful as models for studying the *in vivo* activities of Rad1 and modulators of Rad1.

The invention also provides purified and isolated mammalian Rad1 polypeptides. Presently preferred is a human Rad1 polypeptide (hRad1) comprising the amino acid sequence set out in SEQ ID NO: 2. A presently preferred mouse Rad1 polypeptide (mRad1) is set out in SEQ ID NO: 4. Rad1 polypeptides of the invention may be isolated from natural cell sources or may be chemically synthesized, but are preferably produced by recombinant procedures involving host cells of the invention. Use of mammalian host cells is expected to provide for such post-translational modifications

(*e.g.*, glycosylation, truncation, lipidation, and phosphorylation) as may be needed to confer optimal biological activity on recombinant expression products of the invention. Rad1 products of the invention may be full length polypeptides, biologically active fragments, or variants thereof which retain specific Rad1 biological or immunological activity. Variants, or mutants, may comprise Rad1 polypeptide analogs wherein one or more of the specified (*i.e.*, naturally encoded) amino acids is deleted or replaced or wherein one or more non-specified amino acids are added: (1) without loss of one or more of the biological activities or immunological characteristics specific for Rad1; or (2) with specific disablement of a particular biological activity of Rad1. Particularly preferred mutants of the invention comprise Rad1 polypeptides wherein DNA binding, protein binding or nuclease activities have been modified. Particularly preferred fragments of the invention include highly conserved regions found in mammalian Rad1, *S. pombe* Rad1, and *U. maydis* REC1, corresponding to amino acid residues 57 through 65, 112 through 124, 166 through 171, and 243 through 256 of hRad1.

Variant products of the invention include mature Rad1 products, *i.e.*, Rad1 products wherein leader or signal sequences are removed, having additional amino terminal residues. Rad1 products having an additional methionine residue at position -1 (Met^{-1}-Rad1) are contemplated, as are Rad1 products having additional methionine and lysine residues at positions -2 and -1 (Met^{-2}-Lys^{-1}-Rad1). Variants of these types are particularly useful for recombinant protein production in bacterial cell types.

The invention also embraces Rad1 variants having additional amino acid residues which result from use of specific expression systems. For example, use of commercially available vectors that express a desired polypeptide such as a glutathione-S-transferase (GST) fusion product provide the desired polypeptide having an additional amino terminal residues which result from cleavage at a designed site between the GST component and the desired polypeptide. Variants which result from expression in other vector systems are also contemplated.

The invention further embraces Rad1 products modified to include one or more water soluble polymer attachments. Particularly preferred are Rad1 products covalently modified with polyethylene glycol (PEG) subunits. Water soluble polymers may be bonded at specific positions, for example at the amino terminus of the Rad1 products, or randomly attached to one or more side chains of the polypeptide.

Also comprehended by the present invention are antibodies (*e.g.*, monoclonal and polyclonal antibodies, single chain antibodies, chimeric antibodies, CDR-grafted antibodies and the like) and other binding proteins specific for Rad1 products or fragments thereof. The term "specific for" indicates that the variable regions of the antibodies of the invention recognize and bind Rad1 polypeptides exclusively or with high selectivity (*i.e.*, able to distinguish distinct Rad1 polypeptides from the family of Rad1 polypeptides despite sequence identity, homology, or similarity found in the family of polypeptides), but may also interact with other proteins (for example, *S. aureus* protein A or other antibodies in ELISA techniques) through interactions with sequences outside the variable region of the antibodies, and in particular, in the constant region of the molecule. Screening assays to determine binding specificity of an antibody of the invention are well known and routinely practiced in the art. For a comprehensive discussion of such assays, see Harlow *et al.* (eds), Antibodies; A Laboratory Manual; Cold Spring Harbor Laboratory; Cold Spring Harbor , NY (1988), Chapter 6. Antibodies that recognize and bind fragments of the Rad1 polypeptides of the invention are also contemplated, provided that the antibodies are first and foremost specific for, as defined above, Rad1 polypeptides. As with antibodies that are specific for full length Rad1 polypeptides, antibodies of the invention that recognize Rad1 fragments are those which can distinguish distinct Rad1 polypeptides from the family of Rad1 polypeptides despite inherent sequence identity, homology, or similarity found in the family of proteins. Two such antibody producing hybridomas are the subject of deposits made with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110 in accordance with U.S. Patent & Trademark Office and Budapest Treaty requirements. Hybridomas 267B and 267H were deposited on March 26, 1999 (Accession numbers to be assigned.)

Specific binding proteins can be developed using isolated or recombinant Rad1 products, Rad1 variants, or cells expressing such products. Binding proteins are useful for purifying Rad1 products and detection or quantification of Rad1 products in fluid and tissue samples using known immunological procedures. Binding proteins are also manifestly useful in modulating (*i.e.*, blocking, inhibiting or stimulating) biological activities of Rad1, especially those activities involved in signal transduction. Anti-idiotypic antibodies specific for anti-Rad1 antibodies are also contemplated.

The scientific value of the information contributed through the disclosures

of DNA and amino acid sequences of the present invention is manifest. As one series of examples, knowledge of the sequence of a cDNA for Rad1 makes possible through use of Southern hybridization or polymerase chain reaction (PCR) the identification of genomic DNA sequences encoding Rad1 and Rad1 expression control regulatory sequences such as promoters, operators, enhancers, repressors, and the like. DNA/DNA hybridization procedures carried out with DNA sequences of the invention under moderately to highly stringent conditions are likewise expected to allow the isolation of DNAs encoding allelic variants of Rad1; allelic variants are known in the art to include structurally related proteins sharing one or more of the biochemical and/or immunological properties specific to Rad1. Similarly, non-human species genes encoding proteins homologous to Rad1 can also be identified by Southern and/or PCR analysis and useful in animal models for Rad1-related disorders. As an alternative, complementation studies can be useful for identifying other human Rad1 products as well as non-human proteins, and DNAs encoding the proteins, sharing one or more biological properties of Rad1.

Polynucleotides of the invention are also useful in hybridization assays to detect the capacity of cells to express Rad1. Polynucleotides of the invention may also be the basis for diagnostic methods useful for identifying a genetic alteration(s) in a Rad1 locus that underlies a disease state or states, including cancer (*i.e.*, bladder, head and neck, cancers as well as small cell lung tumors), immune and proliferative disorders, cirrhosis, and rheumatoid arthritis.

Also made available by the invention are antisense polynucleotides which recognize and hybridize to polynucleotides encoding Rad1. Full length and fragment antisense polynucleotides are provided. The worker of ordinary skill will appreciate that fragment anti-sense molecules of the invention include (i) those which specifically recognize and hybridize to Rad1 RNA (as determined by sequence comparison of DNA encoding Rad1 to DNA encoding other known molecules) as well as (ii) those which recognize and hybridize to RNA encoding other members of the Rad1 family of proteins. Antisense polynucleotides that hybridize to multiple DNA encoding other members of the Rad1 family of proteins are also identifiable through sequence comparison to identify characteristic, or signature, sequences for the family of molecules. Antisense polynucleotides are particularly relevant to regulating expression of Rad1 by those cells expressing Rad1 mRNA.

The DNA and amino acid sequence information provided by the present invention also makes possible the systematic analysis of the structure and function of Rad1s. DNA and amino acid sequence information for Rad1 also permits identification of binding partner compounds with which a Rad1 polypeptide or polynucleotide will interact. Agents that modulate (*i.e.*, increase, decrease, or block) Rad1 activity or expression may be identified by incubating a putative modulator with a Rad1 polypeptide or polynucleotide and determining the effect of the putative modulator on Rad1 activity or expression. The selectivity of a compound that modulates the activity of the Rad1 can be evaluated by comparing its binding activity on Rad1 to its activity on other Rad1 enzymes. Cell based methods, such as di-hybrid assays to identify DNAs encoding binding compounds and split hybrid assays to identify inhibitors of Rad1 polypeptide interaction with a known binding polypeptide, as well as *in vitro* methods, including assays wherein a Rad1 polypeptide, Rad1 polynucleotide, or a binding partner are immobilized, and solution assays are contemplated by the invention.

Selective modulators may include, for example, antibodies and other proteins or peptides which specifically bind to a Rad1 polypeptide or a Rad1-encoding nucleic acid, oligonucleotides which specifically bind to a Rad1 polypeptide or a Rad1 gene sequence, and other non-peptide compounds (*e.g.*, isolated or synthetic organic and inorganic molecules) which specifically react with a Rad1 polypeptide or underlying nucleic acid. Mutant Rad1 polypeptides which affect the enzymatic activity or cellular localization of the wild-type Rad1 polypeptides are also contemplated by the invention. Preferred mutants of the invention include those which result in loss of DNA-binding or protein-protein interaction activity, as well as mutants wherein modifications are effected in regions found to be conserved in mammalian and yeast Rad1 polypeptides as discussed above. Presently preferred targets for the development of selective modulators include, for example: (1) regions of the Rad1 polypeptide which contact other proteins and/or localize the Rad1 polypeptide within a cell, (2) regions of the Rad1 polypeptide which bind and/or hydrolyze substrate, (3) allosteric cyclic nucleotide-binding site(s) of the Rad1 polypeptide, (4) phosphorylation site(s) of the Rad1 polypeptide, (5) regions of the Rad1 polypeptide which are involved in multimerization of Rad1 subunits, (6) regions of the Rad1 polypeptide which contact polynucleotides, and (7) regions of the Rad1 polypeptide which possess enzymatic activity. Still other selective modulators include those that

recognize specific Rad1 encoding and regulatory polynucleotide sequences. Modulators of Rad1 activity may be therapeutically useful in treatment of a wide range of diseases and physiological conditions in which Rad1 activity is known, believed, or proposed to be involved as discussed herein.

The present invention further embraces screening assays to identify modulators of Rad1 DNA binding, protein binding and exonuclease activities. Assays of the invention include those comprising the steps of contacting Rad1 and a target ligand (polynucleotide or protein) in the presence and absence of a compound, determining binding between Rad1 and the target ligand in the presence and absence of the compound, identifying the compound as a modulator of Rad1/ligand binding wherein a difference in Rad1/ligand binding is observed in the presence of the compound. In assays of the invention, detecting increased binding in the presence of the compound indicates the compound is an activator and instances wherein decreased Rad1/ligand binding is detected, the compound is identified as an inhibitor. Assays of the invention include those wherein (i) one of Rad1 and the target ligand is immobilized, (ii) the other is detectably labeled, (iii) the ligand and Rad1 are contacted in the presence and absence of a compound, and (iv) the ability of the compound to modulate Rad1/ligand binding is assessed by a change in the amount of immobilized label compared to the amount of immobilized label in the absence of the compound. The invention further embraces solution assays wherein either binding partner Rad1 and/or the target ligand is detectably labeled, the binding partners are contacted in the presence and absence of a compound, binding between Rad1 and the ligand is detected in the presence and absence of the compound, and the compound is identified as a modulator of binding if a change in Rad1/ligand binding is detected.

The invention contemplates that mutations in the Rad1 gene that result in loss of normal function of the Rad1 gene product underlie human disease states in which failure of the cell cycle checkpoint is involved. Gene therapy to restore Rad1 activity would thus be indicated in treating those disease states (for example, various forms of cancer described herein). Delivery of a functional Rad1 gene to appropriate cells is effected *in vivo* or *ex vivo* by use of vectors, and more particularly viral vectors (*e.g.*, adenovirus, adeno-associated virus, or a retrovirus), or *ex vivo* by use of physical DNA transfer methods (*e.g.*, liposomes or chemical treatments). For reviews of gene therapy

technology see Friedmann, *Science, 244*: 1275-1281 (1989); Verma, *Scientific American*, 68-84 (1990); and Miller, *Nature, 357*: 455-460 (1992). Alternatively, it is contemplated that in other human disease states preventing the expression of or inhibiting the activity of Rad1 will be useful in treating the disease states. It is contemplated that antisense therapy or gene therapy could be applied to negatively regulate the expression of Rad1. Antisense nucleic acids (preferably 10 to 20 base pair oligonucleotides) capable of specifically binding to Rad1 expression control sequences or Rad1 RNA are introduced into cells (*e.g.*, by a viral vector or colloidal dispersion system such as a liposome). The antisense nucleic acid binds to the Rad1 target sequence in the cell and prevents transcription or translation of the target sequence. Phosphothioate and methylphosphate antisense oligonucleotides are specifically contemplated for therapeutic use by the invention. The antisense oligonucleotides may be further modified by poly-L-lysine, transferrin polylysine, or cholesterol moieties at their 5′ end.

Moreover, for example, if a particular form of cancer results from a mutation in a gene other than Rad1, an agent which inhibits the transcription or the biological activity of Rad1 and thus the cell cycle checkpoint may be used to render cancerous cells more sensitive to chemotherapy or radiation therapy. The therapeutic value of such an agent lies in the fact that current radiation therapy or chemotherapy in most cases does nothing to overcome the ability of the cancerous cell to sense and correct the DNA damage imposed as a result of the treatment. As a result, a cancer cell can simply repair the DNA damage. Modulating agents of the invention may therefore be chemotherapy and radiation adjuvants or may be directly active as chemotherapeutic drugs themselves.

DETAILED DESCRIPTION

The present invention is illustrated by the following examples. Example 1 details the identification and sequence analysis of the cDNAs encoding the human and mouse Rad1 proteins. Example 2 describes the tissue expression pattern of human Rad1 (hRad1) mRNA. Example 3 details the expression and purification of a recombinant hRad1 fusion protein from bacteria. Example 4 describes the immunological reagents generated for Rad1 detection and purification. The chromosomal localization of the hRad1 gene is described in Example 5. Isolation of the mouse genomic Rad1 (mRad1) gene is

explained in Example 6. Expression of recombinant hRad1 in mammalian cell lines, isolation of stably transfected cell lines, and cellular localization of a fluorescent hRad1 fusion protein is described in Example 7. Example 8 details two biochemical assays for the analysis of hRad1 function. Example 9 describes indirect immunofluorescence assays that localize hRad1 protein within cells undergoing meiosis. Example 10 describes experiments designed to identify proteins that associate with Rad1.

Throughout this Detailed Description the small scale plasmid preparations (minipreps) were a carried out with Wizard Plus kit (Promega), large scale plasmid preparations were done with Midi-Prep or Maxi-Prep kits (Qiagen), DNA purification from electrophoresis gels was performed using Gene Clean II kit (Bio101), all restriction digests and DNA ligation reactions were carried out according to manufacturer's guidelines, and all oligonucleotide synthesis and DNA sequencing was performed using standard procedures routinely practiced in the art.

Example 1
Identification of a Human Rad1 cDNA

In an attempt to identify a human homolog of *S. pombe* Rad1, the polynucleotide sequence of the yeast was utilized as the query sequence in a search of the National Center for Biotechnology Information (NCBI) expressed sequence tag (EST) database. The EST database provides 5′ and/or 3′ nucleotide sequences for cDNA clones from various libraries, and for each EST, the cDNA clone source is also identified. The TBLASTn program available from NCBI was used for nucleotide sequence comparison and to identify ESTs encoding polypeptides having significant homology to the yeast protein sequence.

Three human cDNA clones (470124, 79425, and 667461) and three mouse cDNAs (789687, 790054, and 789468) were identified as being homologous to the yeast gene. Of the human cDNAs, clone 470124 was identified by 5′ EST AA029854 (SEQ ID NO: 5) and 3′ EST AA029300 (SEQ ID NO: 6), clone 79425 was identified by 5′ EST T60007 (SEQ ID NO: 7) and 3′ EST T59937 (SEQ ID NO: 8), and clone 667461 was identified by 5′ EST AA227739 (SEQ ID NO: 9) and 3′ EST AA228124 (SEQ ID NO: 10). Of the mouse cDNAs, clone 789687 was identified by 5′ EST AA387463 (SEQ ID NO: 11), clone 790054 was identified by 5′ EST AA387891 (SEQ ID NO: 12), and clone

789468 was identified by 5'EST AA386981 (SEQ ID NO: 13). All six cDNAs were purchased (I.M.A.G.E.) as *E. coli* strains harboring the cDNA-containing plasmids, and the plasmid DNAs were purified. The three human cDNA clones were sequenced in entirety on both strands using primers that hybridized to the vector DNA and primers designed to hybridize to the human cDNA. Sequences for clones 470124, 79425, and 667461 are set out in SEQ ID NOs: 14, 15, and 16, respectively. Since all three mouse clones had the same 5' sequence as identified in the EST database and restriction analysis (*Eco*RI/*Not*1) revealed that each clone had the same size insert (approximately 1.2 kb), they were predicted to have the same cDNA insertion. Therefore, only one mouse clone (789687, SEQ ID NO: 17) was chosen for complete sequence determination. The predicted polynucleotide and amino acid sequences for hRad1 are set out in SEQ ID NOs: 1 and 2, respectively. The nucleotide and amino acid sequences for mRad1 are set out in SEQ ID NOs: 3 and 4, respectively.

The deduced peptide sequences of the encoded human and mouse proteins were 90% identical to each other over a 280 amino acid sequence which comprised the entire length of the mouse protein. The human protein was two amino acids longer at the carboxy terminus. Comparison of the mammalian proteins with known proteins using BLASTp default search parameters revealed that they were most related to *S. pombe* Rad1 and *U. maydis* REC1. *S. pombe* Rad1 was approximately 32% identical to the human and mouse proteins at three non-overlapping regions that average 70 amino acids in length. *U. maydis* REC1 is approximately 39% identical to human and mouse proteins at five non-overlapping regions that average 37 amino acids in length. Because of the high degree of homology, the human and mouse clones were presumed to be species homologs of the yeast proteins and therefore designated hRad1 and mRad1.

The hRad1 gene was determined to comprise an open reading frame (ORF) of 846 base pairs beginning at the second ATG. Clone 667461 contained the entire ORF. In clone 470124, a 109 base pair deletion (as compared to the ORF in 667461) beginning after nucleotide 293, and correspond to nucleotides 375 to 483 in 667461.

In a subsequent search of the EST database, at least two other cDNA clones were identified that had the same 109 bp deletion. Removal of this 109 bp sequence from the mRNA may indicate alternative splicing (*i.e.*, selection of a 3' splice site 109 bp downstream of the correct one). In fact, DNA sequences preceding the 109

bp sequence (AG) and at the end of the 109 bp sequence (CAG) showed high resemblance to the consensus sequence for exon/intron/exon junctions [Lewin, GENES IV, Oxford University Press:New York (1997) p.88]. The most common sequence at the 3′ end of an exon is AG, and at the 3' end of an intron is CAG. In order to determine if an intron is included in the 667461 sequence, PCR analysis of genomic DNA is used to verify this prediction.

The 3′ ends of hRad1 in 470124 and 667461 were six base pairs apart and two other cDNA clones in the EST database were identified that had the exact 3′ end as 667461. Taken together, these observations suggested that the dominant site for polyadenylation of the hRad1 mRNA was approximately 220 nucleotides 3′ of the termination codon. Adding up the size of the ORF, 3′ untranslated region, and 5′ untranslated region (195 base pairs in 470124), it was predicted that the hRad1 mRNA on should be 1260 nucleotides plus the length of the polyA tail. The additional 358 nucleotides found in the 3′ untranslated region of clone 79425 may be an artifact of cDNA cloning or may represent a population of hRad1 mRNA that utilize more distal polyadenylation sites.

The sequences surrounding the start codons of the hRad1 and mRad1 genes were not highly homologous to a consensus eukaryotic translation initiation sites [Kozak, *et al., J. Mol. Biol. 196*:947-950 (1987)]. If it was assumed that the first hRad1 ATG (rather than the second) was the true initiating codon, the region would be even more divergent from the consensus sequence, thereby supporting the belief that the second ATG represented the beginning of the open reading frame. More significantly, the first ATG codon was not in the correct reading frame.

Example 2
Northern Analysis of hRad1 Expression

Multiple tissue and human cell line Northern blots (Clontech) were probed with hRad1 cDNA to determine the expression pattern of hRad1 mRNA. The approximately 1.2 kb *Eco*RI/*Not*I cDNA insert from 470124 was purified from a 1.2% agarose gel, and 100 ng of the purified DNA was radioactively labeled with [α^{32}P]dCTP using the Random Prime labeling method (Boehringer Mannheim) according to manufacturers instructions. Unincorporated nucleotides were removed by gel filtration

and the specific activity of the probe was determined using a BIOSCAN QC4000. Northern blots were probed according to Clontech instructions, and subjected to autoradiography. Transcript sizes were determined by their location in relation to the location of RNA size standards marked on the blots.

The hRad1 mRNA was present in all tissues tested (spleen, thymus, prostate, testis, ovary, small intestine, colon, leukocyte, heart, brain, placenta, lung, liver, skeletal muscle, kidney, and pancreas) which would be expected for a gene with general cell cycle maintenance functions. In all cell lines (HL-60, HeLa S3, K562, MOLT-4, Raji, SW480, A549, and G361) probed hRad1 was also present. Three bands of approximately 1400, 2800, and 5200 nucleotides, however, were detected in all cells. In leukocytes, liver, skeletal muscle, and kidney, an additional weak band of about 1900 nucleotides was observed. The 1400 nucleotide mRNA signal, which closely corresponded to the size of the cDNA identified in Example 1 was of equivalent or higher intensity to the other hybridization signals in each lane. Expression of the 1400 nucleotide transcript was highest in testis and heart, and lowest in leukocyte and lung. In the cell lines, the 1400 nucleotide transcript was most abundant in HeLa S3 and K562, and least abundant in A549 (a lung carcinoma cell line).

The origin of the larger mRNAs detected may indicate alternative processing of the hRad1 transcript or result from a Rad1-related gene family. Alternatively, the probe may have contained sequences that hybridize to transcripts unrelated to Rad1. In view of the different lengths of cDNA clones identified in Example 1, it is possible that the hRad1 3′ untranslated region may be polyadenylated at sites that are at least 350 bp apart which could account for the different transcript sizes. In addition, alternative splicing of hRad1 mRNA may occur. As another possibility, a hRad1 pseudogene (see Example 5), if transcribed in either direction, would likely generate a hybridization signal on the Northern blots. It is noted that Western blot analysis, discussed below, suggests that more than one Rad1 gene product is expressed.

Example 3
Expression of hRad1 in Bacteria

The hRad1 coding sequence was cloned into a bacterial expression vector in order to express glutathione-S-transferase (GST) fusion protein (GST-hRad1). This

strategy was chosen because of the ease of purification of GST proteins on a reduced glutathione (GSH) agarose matrix. Additionally , the fusion protein contains a Factor Xa protease cleavage site between the GST and hRad1 moieties, allowing the excision of hRad1 when required.

Generation of the GST-hRad1 expression plasmid, pGEX-hRad1, was performed as follows. An *Eco*RI/*Not*I fragment of approximately 1.2 kb from the 470124 clone was gel purified and ligated between the *Eco*RI and *Not*I sites of pGEX5x-1 (Pharmacia Biotech) to produce pGEX-470124.EN. A PCR primer (Rad1.ATGPCR, SEQ ID NO: 18) was designed to modify the DNA sequence upstream of the start codon of hRad1 to generate convenient cloning sites (*Eco*RI and *Bam*HI). Rad1.ATGPCR also encoded the first five hRad1 codons at the 3′ end and 18 bp restriction site containing sequence at the 5' end. PCR was performed with Rad1.ATGPCR and Rad1.5 (SEQ ID NO: 19) oligonucleotides to amplify a 994 base pair fragment from 667461 under the following conditions. A 50 µl reaction mixture contained a 667461 colony (transferred with a toothpick), 200 ng Rad1.5, 400 ng Rad1.ATGPCR, 3 units of *Taq* DNA polymerase (AmpliTaq; Perkin Elmer), 1X Perkin Elmer buffer with $MgCl_2$, and 0.2 mM dNTPs. Thirty thermocycles (94°C for one minute, 56°C for 30 seconds, 72°C for 1.5 minutes) were followed by a five minute incubation at 72°C. One tenth of the reaction was analyzed by agarose gel electrophoresis and the remainder was phenol/chloroform extracted, and ethanol precipitated by standard procedures. The PCR product was resuspended in water and digested with *Eco*RI and *Sty*I to generate a fragment of approximately 351 base pairs which was purified from an agarose gel. Vector pGEX-470124.EN was also digested with the same two enzymes and the fragment containing the vector backbone and the 3′ portion of hRad1 was purified from an agarose gel. Both vector and insert *Eco*RI/*Sty*I fragments were ligated together at 14°C under standard conditions (New England Biolabs) and transformed into competent *E. coli* TOP10F' cells (Invitrogen) according to the manufacturer's suggested protocol and transformants were selected on LBM agar containing carbenicillin. Colonies were screened by PCR under the above conditions, with the following exceptions: reaction volumes were 20 µl, primers were Rad1.ATGPCR (200 ng) and Rad1.1 (SEQ ID NO: 20) (100 ng), and thermocycles were 94°C for 30 seconds, 50°C for 30 seconds, and 72°C for 30 seconds. Transformants containing pGEX-hRad1 produced an amplification product of 540 base

pairs.

The GST-hRad1 fusion protein was purified on a glutathione matrix as follows. The GST protein (from pGEX5x-1) was purified concurrently by this same method. Cultures of pGEX-hRad1 (in TOP10F' cells) were grown at 30°C in LBM with 0.1 mg/ml carbenicillin to an absorbance of 0.5 to 0.7 at 600 nm at which point IPTG (1 mM final concentration) was added to induce gene expression. Five hours post-induction the cells were centrifuged at 10,000 x *g* for 10 minutes. Cell pellets were stored at -70°C before use. Thawed pellets were resuspended (0.1 ml per ml x A_{600} units of the culture) in lysis buffer (25 mM Tris-HCl, pH 7.5, 10 mM EDTA, 0.1% NP-40, 150 mM NaCl, 5 mM DTT, 0.5 mM AEBSF, 1 μg/ml leupeptin (a protease inhibitor), 5 μg/ml aprotinin, and 2 μg/ml pepstatin). Cells were disrupted by sonication and centrifuged for 20 minutes at 10,000 x *g*. The supernatant was decanted and the pellet was saved for a different purification protocol (see below). The supernatant was mixed with 1/50th volume of glutathione agarose (Sigma) for 30 minutes at 4°C. The GSH agarose was sedimented by low speed centrifugation and the pellet was washed four times with 33 volumes of wash buffer (10 mM Tris-HCl, pH 8, 1 mM EDTA, 0.1% NP-40, 150 mM NaCl, 0.1 mM AEBSF, 1 μg/ml aprotinin, and 0.4 μg/ml pepstatin). Proteins were eluted from the glutathione matrix with 3.3 volumes of 20 mM glutathione buffer (containing 120 mM NaCl and 50 mM Tris-HCl, pH 9.5). Elution was carried out for ten minutes at 0°C and the slurry was drained through a PolyPrep column (BioRad) to remove the matrix. The eluate was dialyzed at 4°C against storage buffer (20 mM Tris-HCl, pH 8, 0.1 mM EDTA, 10% glycerol, 50 mM KCl, and 1 mM DTT). Protein concentration was determined with Pierce Coomassie Plus with BSA as a standard. Protein was stored in aliquots at -70°C after freezing in liquid nitrogen.

The pellet from the above cell extract also served as a source for purification of GST-hRad1. It was observed from analysis of Coomassie stained protein gels and on Western blots that about 90% of the GST-hRad1 protein produced in these bacteria remained insoluble following sonication. The pellet was resuspended (½ volume of cell extract) in Buffer A (25 mM Tris-HCl, pH 7.5, 1 mM EDTA, 0.1% NP-40, 1 mM DTT, 0.5 mM PMSF, 1 μg/ml leupeptin, and 1 μg/ml aprotinin) and was centrifuged for 20 minutes at 10,000 x *g*. The supernatant was removed and the wash step was repeated. The pellet was resuspended in the same volume of Buffer A but also containing 2 M NaCl

and centrifuged as above. The supernatant was removed and the wash step was repeated. The pellet was resuspended in the same volume of Low Urea Buffer (2.5 M urea and 50 mM Tris-HCl, pH 8) and centrifuged as above. The final pellet was resuspended (1/4 volume of cell extract) in High Urea Buffer (6.65 M urea and 50 mM Tris-HCl, pH 8), incubated at room temperature for 10 minutes with occasional agitation, and centrifuged as above. The resulting supernatant, designated GST-hRad1-sol6, was a highly purified protein preparation. Although the protein concentration of GST-hRad1-sol6 was high (approximately 1 to 2 mg/ml), it could be dialyzed directly into nondenaturing buffer, calcium-, magnesium-free PBS (CMF-PBS), with no significant loss of protein due to insolubility. Aliquots were stored before and after dialysis at -70°C after freezing in liquid nitrogen.

Example 4
Production of hRad1 Antibodies

Recombinant GST-hRad1 and a synthetic peptide were used as antigens to raise antibodies to hRad1 in rabbits and mice.

The hRad1 peptide sequence was provided to Quality Controlled Biochemicals, Inc. (QCB), which was contracted to identify a likely immunogenic peptide sequence, synthesize the peptide, and immunized rabbits for the production of antisera. The synthetic peptide (acetyl-NQTQVNRYKISSLLKPSC-amide) corresponded to amino acids 216-231 of hRad1 (SEQ ID NO: 2) with an additional carboxy terminal cysteine to aid in cross-linking of the peptide.

Anti-peptide sera from two rabbits were shown to recognize GST-hRad1 by Western blot and detection of GST-hRad1 could be prevented by pre-incubating the synthetic peptide with the antisera before probing Western blots.

In order to purify antibodies from the rabbit antisera, the synthetic peptide was covalently cross-linked to a solid matrix (Sulfo-Link; Pierce) according to manufacturer's suggested protocol and the resulting affinity matrix (approximately 44 μg peptide/ml) was used in affinity purification. Sera was diluted two-fold with 50 mM sodium phosphate (pH 6.5) and loaded on the column at 0.25 ml/minute. The column was washed with phosphate buffer containing 0.5 M NaCl and 0.02% sodium azide.

Antibodies were then eluted with 100 mM glycine (pH 2.5) and the pH of eluate fractions was neutralized with 1/10th volume 1 M Tris-HCl (pH 8). Eluates were dialyzed against 0.9X CMF-PBS 10% glycerol and stored at 4°C.

Western blot analysis determined that the affinity-purified antisera recognized GST-Rad1, GFP-Rad1, endogenous HeLa cell proteins of approximately 30 and 35 kDa), and four endogenous proteins from mouse testes extract (each approximately 30, 35, 65, and 95 kDa).

In another approach to antibody production, R & R Rabbitry (Stanwood, WA) was contracted to immunize rabbits with GST-hRad1. Antisera from these rabbits was demonstrated to recognize GST-Rad1 recombinant protein on Western blots. Endogenous proteins from human cell lines (30 and 37 kDa) were also detected in these assays.

In a third approach to antibody production, GST-hRad1 was also used to immunize 4 mice. Mice were initially immunized by subcutaneous injection of 40 μg GST-hRad1 in Complete Freund's Adjuvant and boosted every three weeks with 20 μg GST-hRad1 in Incomplete Freund's Adjuvant. Once sera reactive to recombinant hRad1 in ELISA and Western blot assays was obtained, one mouse was chosen for production of hybridoma cell lines. A pre-fusion boost of 50 μg GST-hRad1 in calcium-, magnesium-free PBS (CMF-PBS) was administered four days prior to the fusion protocol. Hybridoma fusions were generated according to previously published protocols [Harlow, *et al., supra*]. Twelve isolated hybridomas (267A-267L) were cloned and their activity in ELISA, Western and immunoprecipitation assays was evaluated. Supernatant from all clones recognized three bands in human (approximately 30, 37, and 66 kDa) and mouse (approximately 28, 35, and 58) cell extracts on Western blots. All twelve monoclonal antibodies could immunoprecipitate the 37 kDa and 66 kDa bands from human cell extracts (immunoprecipitate of the 30 kDa band was inconclusive due to the migration of IgG light chains in same region). Ascites were made with 267B and 267H cells.

Two of the above described hybridomas are the subject of deposits made with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110, in accordance with U.S. Patent & Trademark Office and Budapest Treaty requirements. Hybridomas 267B and 267H were deposited on March 26, 1999 (Accession numbers to be assigned).

Example 5
Chromosomal Localization of hRad1

Gene-specific oligonucleotides were used in PCR reactions with human, mouse and hamster genomic DNA to find primer sets that would amplify specific human gene fragments. Subsequently, PCR reactions with the NIGMS human/rodent somatic cell hybrid mapping panel #2 (Coriell Cell Repositories) were performed localize the PCR fragments to specific human chromosomes. Fine mapping was performed by PCR with the Stanford G3 Radiation Hybrid (G3/RH) DNA panel (Research Genetics). DNA sequence analysis of the PCR fragments distinguished the Rad1-derived from pseudogene-derived amplification products.

PCR reaction on the entire somatic cell hybrid panel was performed with primers Rad1.ATGPCR (SEQ ID NO: 18) (200 ng) and Rad1.3 (SEQ ID NO: 21) (100 ng) in 25 μl volumes. Buffer conditions for these and other PCR reaction described in this example included 200 μM dNTPs, 1X Perkin Elmer *Taq* polymerase Buffer, 1X Perkin Elmer $MgCl_2$. One microliter (approximately 300 ng) of somatic cell hybrid DNA or control DNA (human, mouse, and hamster) was added in individual reaction mixtures. Additional controls with no template and 10 pg 667461 plasmid DNA were performed. Primers and template were incubated at 94°C for 15 seconds, followed by addition of buffer, dNTP, $MgCl_2$, and 1.25 units *Taq* DNA polymerase (AmpliTaq®; Perkin Elmer). Incubation at 94°C continued for one to eight minutes, followed by 40 thermocycles (94°C for one minute, 50°C for one minute, 72°C for one minute). After a final extension step for five minutes at 72°C, the samples were analyzed by electrophoresis on 2% agarose.

Reactions with three somatic cell hybrid DNAs produced PCR products of equivalent size to those seen with human genomic DNA and 667461 plasmid DNA (135 base pairs), indicating that human chromosomes 5, 6 and 10 contained Rad1-related sequences.

To further characterize the location of the hRad1 gene, different primer sets were used on somatic cell hybrid DNAs (containing chromosomes 5, 6 and 10), 667461 plasmid, human cDNA, and hamster cDNA. Reactions varied from those above as follows. Final reaction volume of 20 μl contained one unit of AmpliTaq® and thermocycle times and temperature varied slightly (94°C for 30 seconds, 52°C for one

minute, 72°C for one minute). PCR product sizes were analyzed on 8% polyacrylamide or 2% agarose gels.

Results indicated that neither Rad1.ATGPCR or Rad1.3 alone directed the amplification of 135 base pair DNA. A combination of Rad1.2 (SEQ ID NO: 22)and Rad1.3 (SEQ ID NO: 21) primers amplified 139 bp products with human genomic, chromosome 5, chromosome 6 and 667461 DNA. Primers Rad1.4 (SEQ ID NO: 23) and Rad1.5 (SEQ ID NO: 19) amplified a 367 bp band from 667461, human genomic, and chromosome 10 DNA, but the most prevalent band in human genomic DNA was approximately 300 base pairs larger (approximately 670 bp) that was also found in chromosome 5 and 6 DNAs. These results indicated that chromosome 5 and 6 contained Rad1-related sequences that included 5′ untranslated region sequence (Rad1.2) and perhaps contained an intron between Rad1.4 and Rad1.5 sequences. Chromosome 10 contained Rad1-related sequences similar in size to the hRad1 cDNA between Rad1.4 and Rad1.5, but did not appear to have 5′ untranslated region sequence of the Rad1.2 primer.

Additional primers (Rad1.7, SEQ ID NO: 24, and Rad1.8, SEQ ID NO: 25) were designed to use in combination with Rad1.4 and Rad1.5 in order to find a primer set that was specific for a single human chromosome. PCR conditions from the previous experiment were modified to include two units of AmpliTaq® and 35 thermocycles which varied according to which primer set was employed. For primer sets Rad1.4 with Rad1.5 and Rad1.4 with Rad1.7, thermocycles were 94°C for 20 seconds, 50°C for 30 seconds, and 72°C for 30 seconds. For primer set Rad1.5 with Rad1.8, thermocycles were 94°C for 20 seconds, 56°C for 30 seconds, and 72°C for 30 seconds. PCR products were analyzed on 8% polyacrylamide gels.

Results indicated that primers Rad1.4 and Rad1.5 amplified the cDNA-size product (367 base pairs) and a larger product (approximately 670 base pairs) from human genomic DNA, the 670 base pair product from chromosome 5 DNA, and no products with chromosome 6 DNA. The chromosome 10 DNA reaction was inconclusive with Rad1.4 and Rad1.5 primers. Rad1.4 and Rad1.7 amplified two bands from human genomic DNA, one approximately 289 base pairs (cDNA size) and the other approximately 600 bp. The Rad1.4 and Rad1.7 amplified the 600 base pair product from chromosome 5 and 6 DNA, and the 289 base pair product from chromosome 10 DNA. These results mirrored results initially observed with Rad1.4 and Rad1.5 PCR reactions.

Rad1.5 and Rad1.8 amplified the expected 331 base pair DNA from 667461 reactions, but not from human genomic or specific human chromosome (5, 6 and 10) DNAs. The primary PCR product observed using primer pair Rad1.5 with Rad1.8 PCR in human genomic, chromosome 5 and chromosome 6 DNAs was approximately 650 base pairs and no product was observed from chromosome 10 reactions.

A radiation hybrid panel created at the Stanford Human Genome Center (G3/RH, purchased from Research Genetics) was screened by PCR with the Rad1.4 and Rad1.7 primer set. Twenty microliter reactions included 100 ng each primer, 2 units AmpliTaq®, 100 ng G3/RH DNA (or 10 pg 667461) in buffer conditions described above. Reaction mixtures were incubated for three minutes at 94°C followed by 35 thermocycles (94°C for 20 seconds, 50°C for 30 seconds, 72°C for 30 seconds) and an additional five minute incubation at 72°C. Products were analyzed on 8% polyacrylamide gels and examined for presence of the 289 base pair (cDNA size) and 600 base pair products. Scores from the 83 G3/RH DNAs were sent to the RH server at Stanford (http://www-shgc.stanford.edu/RH/rhserver_form2.html).

By analyzing the map locations of the closely linked STS DNAs (SHGC-33199 and SHGC-8907) and their flanking markers, it was determined that SHGC-33199 maps to 10q25.3 and SHGC-8907 maps to 5p13.1.

Sequence analysis of the 289 base pair and 600 base pair PCR products (using primer pair Rad1.4 and Rad1.7) was performed to determine which product corresponds to the actual hRad1 gene. PCR reactions were as described for the G3/RH panel above. Products were separated on an 8% polyacrylamide gel and purified by Gene Clean II Kit (Bio101). Three different 289 base pair products were sequenced: 1) combined fragments (considered wild type) from reactions containing human genomic, G3/RH-48, and G3/RH-69 DNA; 2) fragment from reactions containing U251 cell line DNA (glioblastoma cell line with possible mutation at 10q25.3) [Perghouse, *et al., Cancer Res. 53*:5043-5050 (1993)]; and, 3) combined fragments from reactions containing DNA from cell lines U251.N10.3, U251.N10.4, and U251.N10.5.6 (U251 cell lines with partial phenotype reversal by insertion human chromosome 10 DNA) [Perghouse, *et al., supra*]. Purification of genomic DNA was carried out as follows.

Genomic DNA was purified from glioblastoma cell lines (U251, U251.N10.3, U2510.N10.4, and U251.N10.5.6) (University of Texas, M.D. Anderson).

The cells were grown in DMEM-F12 including 5% fetal bovine serum to near confluence on 100 mm plates. After aspirating the media, cells were washed twice with 10 ml CMF-PBS, then were mechanically dislodged in 5 ml CMF-PBS. Cells were centrifuged for four minutes at 2000 x *g*, resuspended in 1 ml CMF-PBS, and centrifuged again. The cell pellets were frozen in liquid nitrogen and stored at -70°C. Cell pellets were thawed on ice, resuspended in 0.2 ml TE (10 mm Tris-HCl, 1 mM EDTA, pH 8), and 0.5 ml lysis buffer (0.6% SDS, 10 mM Tris-HCl, pH 7.5, 10 mM EDTA, 0.2 mg/ml Proteinase K, 20 μg/ml RNase A) was added slowly. The suspension was incubated 30 minutes at 37°C. The aqueous mixture was phenol extracted three times and extracted one time with chloroform/isoamyl alcohol (24:1). DNA was precipitated with 1/10th volume sodium acetate (pH 5.2) and 2.5 volumes of ethanol and centrifuged immediately for 3 minutes at 3000 rpm in an Eppendorf microfuge. The DNA pellets were resuspended in TE and concentration was determined by measuring absorbance at 260 nm.

The 600 base pair product was sequenced from combined fragments from reactions containing U251, U251.N10.3, U251.N10.4 and U251.N10.5.6 DNAs. Primers for sequencing were the same as used for PCR. All three 289 base pair fragment preparations contained the same DNA sequence which differed from the cDNA at 12 nucleotides (resulting in seven missense mutations) and included a five nucleotide deletion causing a frameshift mutation indicating that the product was derived from a pseudogene. The 600 base pair fragment matched the cDNA sequence except for a 312 nucleotide region that was a probable intron based on the presence of conserved sequences at the flanking ends.

Intron placement was after nucleotide 665 in the open reading frame (nucleotide 841 from the 5′ end of the 667461 sequence). Because of the presence of this intron and the sequence flanking this region, it is possible that the 600 nucleotide Rad1.4/Rad1.7 PCR product was from the *bona fide* Rad1 gene (which mapped to 5p13.1). The Rad1-related sequence on chromosome 10q25.3 was most likely a pseudogene since the open reading frame was interrupted and may not have been transcribed since no sequences in EST database were found that matched that of the 289 nucleotide sequence.

Several tumor cell lines have been reported that have mutations in the 5p13.1 region [Carey, *et al.*, *Anticancer Res. 13*:2561-2568 (1993); Van Dyke, *et al.*,

Genes, Chromosomes and Cancer 9:192-206 (1994); Bohn, *et al., Intl. J. Cancer 74*:291-295 (1997), Levin, *et al., Genes, Chromosomes and Cancer 13*:175-185 (1995)] suggesting that disruption or amplification of the Rad1 gene may contribute to the neoplastic phenotype.

Example 6
Isolation of Mouse Genomic Rad1 DNA

In order to isolate genomic DNA encoding mouse Rad1, a PCR screening method was developed as follows. PCR mixtures (20 μl total volume) contained 1X Perkin Elmer buffer with $MgCl_2$, 200 μM dNTPs, 100 ng each of muR1.3 (SEQ ID NO; 26) and muR1.5 (SEQ ID NO: 27) primers, 308 ng mouse genomic DNA (Coriell Cell Repositories) or 10 pg 789687 DNA, and 2 units of AmpliTaq® polymerase. Reactions mixtures were heated at 94°C for three minutes, followed by 40 thermocycles (94°C for 20 seconds, 55°C for 30 seconds, 72°C for 30 seconds), and a final incubation at 72°C for five minutes.

A 127 base pair genomic product was amplified and detected on a 10% polyacrylamide gel. The same size product was amplified from mRad1 cDNA (789687 plasmid). These amplification conditions and the primers (muR1.3 and muR1.5) were supplied to Genome Systems to screen a mouse genomic library in phage P1 (PI-4535; mouse ES 129/OLA).

Three P1 clones (P1-16393, -16394, -16395) were identified in this screen at Genome Systems. These clones (phage and purified DNA), predicted to contain 75-85 kb of mouse DNA, were analyzed as follows. Using the above PCR reaction conditions (with 100 ng P1-clone DNA) with two different primer sets (muR1.3 with muR1.5 and muR1.1, SEQ ID NO: 28, with muR1.4 SEQ ID NO: 29), one of the three P1 clones (P1-16394) was found to be positive for both expected amplification products. The P1-16394 PCR products were excised from the 10% polyacrylamide gel and purified with Gene Clean II Kit (Bio101). The purified fragments were sequenced, which revealed that both P1-16394 PCR products matched the mRad1 cDNA sequence. P1-16394 was selected to subclone the genomic copy of the mRad1 gene.

Example 7
Expression of hRad1 in Mammalian Cells

In order to obtain inducible expression of the human Rad1 gene in mammalian cells, the following protocol was carried out.

Construction of Expression Vectors

Expression vectors were designed to express native hRad1 and two epitope-tagged hRad1 fusion proteins (GFP-hRad1 and FH-hRad1). GFP-hRad1 was designed to express the hRad1 protein as a fusion protein with the Green Lantern Green Fluorescent Protein (GFP) which allows microscopic detection of the fluorescent protein within the cell under ultraviolet light. FH-hRad1 was designed to express hRad1 protein as a fusion in combination with a FLAG epitope (for immuno-detection and immuno-purification) and an epitope consisting of six consecutive histidines for immunodetection and affinity purification on nickel-agarose. Vector construction and analysis of transfected cells is described below.

Construction of pIND-hRad1, a vector for inducible expression of hRad1, was performed as follows. The hRad1 gene was excised from pGEX-hRad1 (see Example 3) digested with *Bam*HI and *Not*I to provide a 1.2 kb gene fragment. The *Bam*HI/*Not*I fragment was purified from a 0.6% agarose gel using Gene Clean II (Bio 101). The vector pIND (Invitrogen) was similarly digested and purified. The pIND vector and hRad1 insert DNAs were ligated together by standard protocol and transformed into *E. coli* TOP10F' cells (Invitrogen). Single colonies were assayed by PCR (as in Example 1) with Rad1.1 and Rad1.ATGPCR primers to detect clones that had the hRad1 insert. DNA from six positive pIND-hRad1 clones was prepared and analyzed by restriction enzyme digestion (*Bam*HI alone, *Not*I alone, and *Bam*HI/*Not*I). Sequence of the 5' junction was verified in DNA sequencing reactions.

Construction of pIND-GFP-hRad1, a vector for inducible expression of GFP-hRad1, was performed as follows. Plasmid pIND-βG-GLGFP contains the β globin intron (to increase translation efficiency) inserted between the *Nhe*I and *Hin*dIII sites of pIND, followed by the Green Lantern green fluorescent protein (GLGFP) gene inserted between the *Hin*dIII and *Bam*HI sites of pIND. The pIND-BG-GLGFP vector was digested with *Bam*HI and *Xho*I and gel-purified as described above. The linear vector and

the hRad1 (*Bam*HI/*Not*I) fragment from pGEX-hRad1 were ligated together in the presence of *Xho*I/*Not*I linker oligonucleotides (SEQ ID NOs: 30 and 31). Ligation mixes were transformed into *E. coli* TOP10F' cells (Invitrogen) and transformants were analyzed by PCR as described above. DNA from five clones was analyzed by restriction digestion (*Bam*H1 alone, *Xho*I alone, and *Bam*HI/*Xho*I) and sequence of the 5' junction was verified in DNA sequencing reactions to ensure that hRad1 would be translated in the same reading frame as the upstream GFP gene.

Construction of pIND-FH-hRad1, a vector for inducible expression of FH-hRad1, was performed as follows. A gene cassette that encoded the FLAG® epitope tag (Asp-Tyr-Lys-Asp-Asp-Asp-Asp-Lys) (SEQ ID NO: 32) and six histidines (FH) was originally cloned into pBAR an arabinose-inducible *E. coli* vector, to produce pBAR8A. A 128 bp *Xba*I/*Bam*HI fragment from pBAR8A was gel-purified and a ligation reaction was performed with the following gel-purified DNA fragments: the 128 bp *Xba*I/*Bam*HI FLAG-6His cassette, the *Bam*H1/*Xho*I hRad1 gene from pIND-GFP-hRad1, and pIND previously digested with *Nhe*I and *Xho*I. The ligation mix was transformed into *E.coli* TOP10F' cells (Invitrogen) and transformants were analyzed by restriction digest of purified plasmid DNA (*Xho*I/*Not*I and *Bam*HI/*Xho*I). One clone (pIND-FH-hRad1) was chosen for sequence analysis of 5' and 3' junctions.

hRad1 Expression

Transient transfection assays of COS cells was used as an initial screen for the inducible expression of hRad1 and GFP-hRad1. The plasmid pVgRXR (Invitrogen) encoding an ecdysone receptor, was co-transfected with pIND-hRad1 and pIND-GFP-hRad1 to generate COS cell transfectants that inducibly express Rad1 or GFP-Rad1 protein in the presence of muristerone. COS cells were grown to approximately 50% confluence in DMEM plus 10% FBS (containing penicillin and streptomycin) in 100 mm plates. Five micrograms of pVgRXR and pIND-hRad1 (or pIND-GFP-Rad1) were co-transfected with Superfect (QIAGEN) reagent, according to the manufacturer's protocol (two co-transfected plates per plasmid combination). After 20 hours, the media was replaced on half the plates with media containing 1 μM muristerone to induce gene expression. At twenty-one hours post-induction, cells were monitored under UV light for GFP fluorescence. GFP-Rad1 expression was observed in induced cells and its cellular

localization was both the nuclear and cytoplasmic. At 24 hours post-induction, cells from transfection plates (induced and uninduced) were harvested (washed two times with 10 ml CMF-PBS, dislodged in 5 ml CMF-PBS, centrifuged four minutes at 1000 rpm, frozen in liquid nitrogen and stored at -70°C). Cell extracts were prepared from thawed cell pellets by resuspension in NP-40 Lysis Buffer (2 mM DTT, 1% NP-40, 50 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EGTA, 10 mM NaF, 1 mM PMSF, 1 μg/ml pepstatin, 1 μg/ml aprotinin, and 1 μg/ml leupeptin). The suspension was incubated for 30 minutes at 4°C with rocking. Particulate matter was removed by centrifugation for five minutes in an Eppendorf microfuge. Supernatants were transferred to a clean tube and protein concentration was determined with Coomassie Plus (Pierce) using BSA as a standard. Western blot analysis with anti-hRad1 antibodies (see Example 4) and anti-GFP antisera (Clontech) confirmed that GFP-Rad1 was inducibly expressed. Expression of hRad1 from pIND-hRad1 was not conclusive.

Stable pIND-GFP-Rad1 and pIND-FH-Rad1 transfectants in ECR293 cells were produced as follows. ECR293 cells (Invitrogen) are a stably transfected cell line containing pVgRXR. ECR293 cells were grown in DMEM plus 10% FBS (with penicillin and streptomycin) and 400 μg/ml zeocin to insure the presence of pVgRXR in the genome. Cells were transfected with 5 μg plasmid DNA (pIND-FH-Rad1 or pIND-GFP-hRad1) with Superfect Reagent (QIAGEN). Stable transfectants were selected in the presence of 400 μg/ml G418 after a 1:10 split onto 100 mm plates. Media was replaced every two to three days. After 18 days of growth (and death of untransfected cells) in G418, 20 growth foci were isolated and transferred to 96 well dishes. After expansion of clones, cells were analyzed for induction of FH-Rad1 and GFP-Rad1 proteins (induction as described for COS cells). Fluorescent GFP-Rad1 was present throughout the cells. Stable transfectants were maintained in media containing 200 μg/ml zeocin and G418, and could be stored in 10% DMSO 90% media in liquid nitrogen tanks.

Cell extracts were prepared by resuspending ECR293 transfected cells (harvested from a 60 mm plate) in 0.2 ml Hypotonic Lysis Buffer (20 mM Tris-acetate, pH 8, 1.5 mM $MgCl_2$, 5 mM KCl, 1 mM DTT, 0.5 mM AEBSF, 2 μg/ml leupeptin, and 2 μg/ml aprotinin) and incubating on ice for 15 minutes, followed by sonication for 30 seconds (setting 7 with cup-horn attachment). NaCl and EDTA were added to a final concentration of 250 mM and 2 mM, respectively. After 30 minute incubation on ice, the

particles were removed by centrifugation for five minutes in an Eppendorf microfuge. Protein concentrations of the cell extracts were determined by Coomassie Plus reagent (Pierce) with BSA as the standard.

Western blot analysis of cell extracts (35 μg/lane) was used to detect the proteins. Rad1 monoclonal antibodies and Rad1 rabbit antisera detected both 35 kDa FH-Rad1 and 56 kDa GFP-Rad1, as well as endogenous proteins (see Example 4). GFP-hRad1 was also detected with GFP antisera (Clontech). FH-hRad was also detected with a FLAG monoclonal antibody (Eastman Kodak) and a 6His monoclonal antibody (Clontech). In the absence of induction, expression of FH-Rad1 and GFP-Rad1 was barely detectable while induction of FH-HRad1 and GFP-hRad1 resulted in protein levels ten times greater than endogenous Rad1 protein levels. Preliminary results indicated that over-expression of GFP-Rad1 or FH-Rad1 in ECR293 cells led to down regulation of the putative endogenous Rad1 protein, but not of the cross-reacting 37 and 66 kDa proteins. Expression levels (24 hr induction) did not increase by increasing muristerone in the media above 1 μM levels (5 or 20 μM), or increasing the induction times from 24 to 48 hours. (see Example 10).

Precipitation assays with induced FH-Rad1 cell extract demonstrated that FH-hRad1 could be precipitated with all 12 Rad1 monoclonal antibodies, FLAG® monoclonal antibody, and 6His monoclonal (plus protein G Sepharose; Pharmacia) and nickel agarose (QIAGEN). Nickel agarose and α FLAG®/monoclonal antibody column matrices are used in the large scale purification of FH-Rad1.

Example 8
Nuclease and DNA Binding Activity of Rad1

Determination of Nuclease Activity

Biochemical assays were performed with GST-hRad1 and two HeLa Rad1 complexes in order to test for nuclease and DNA-binding activity based on the previous reports that the Rad1 homolog, *U. maydis* REC1, expressed and purified from bacteria has been shown to have 3' to 5' exonuclease activity [Onel, *et al., Mol. Cell. Biol. 15*:5329-5338 (1995); Thelen, *et al., J. Biol. Chem. 269*:747-754 (1994)]. No *in vitro* activity for *S. pombe* Rad1 has been reported.

The assay for hRad1 exonuclease activity was performed essentially as

previously described for *U. maydis* REC1 assays [Thelen, *et al., supra*]. GST-Rad1 was purified on glutathione agarose and dialyzed against storage buffer as described in Example 3. HeLa Rad1 complexes were purified as described in Example 10. Exonuclease substrate was *Sau*3AI digested pGEX-5X-1 DNA that was 3' end labeled with Klenow polymerase in the presence of 2 μM [α-^{32}P]dCTP (~10^6 cpm/pmol) and 25 μm each of dATP, dGTP, and dTTP. Alternatively, 3′ end labeled pAdUAS (3 kbp plasmid linearized at single NdeI site) served as the substrate. Unincorporated label was removed by purifying DNA using Gene Clean II Kit (Bio 101) or by using Micro Bio spin columns (BioRad). Standard reaction conditions included 20 mM Tris-HCl (pH 8), 10 mM $MgCl_2$, 1 mM DTT, and 1-30 μM (total nucleotide)substrate. Variables tested included (i) pH ranges from 6.8 to 9.5, (ii) use of heat denatured DNA, (iii) adding 1 mM $MnCl_2$ and/or 100 μM zinc sulfate, (iv) adding 30 mM NaCl, (v) adding 0.05% NP-40, and (vi) adding 100 µg/ml BSA. Reaction mixtures (generally 20-200 μl) were incubated at 37°C and stopped with a minimum of 1.5 volumes of 0.5 mg/ml sheared and denatured fish DNA (Boehringer Mannheim) in 25 mM EDTA. The stopped reactions were placed on ice and an equal volume of ice cold 10% trichloroacetic acid was added. Incubation on ice was continued for 10 minutes before centrifugation for 10 minutes at 4°C (approximately 15,000 x *g*). The supernatant was removed to scintillation vials containing 5 ml EcoScint (National Diagnostics, Atlanta, GA) and radioactivity was determined on a scintillation counter. In this assay unhydrolyzed DNA fragments are TCA-precipitated and free nucleotides (released by DNA nucleases) are TCA-soluble.

Results indicated that the GST-hRad1 protein preparation releases labeled nucleotides from the 3′ end of DNA. The nuclease activity of GST-hRad1 was linear from 27 to 159 minutes. Heat-denatured DNA served as a better substrate in these reactions. Both NP-40 and BSA had slight stimulatory effects on activity, while $ZnSO_4$ and $MnCl_2$ were inhibitory. Sodium chloride (30 mM) did not affect GST-hRad1 nuclease activity. Exonuclease activity directly correlated with amount of GST-hRad1 present in the reaction (2 to 9.2 μg/ml were tested). GST protein, a negative control in these assays, had detectable but much lower nuclease activity than GST-hRad1. Additionally, the larger of the two Rad1 complexes purified from HeLa cells has a co-purifying nuclease activity that is strongly biased toward denatured DNA substrates. The small Rad1 complex lacked significant nuclease activity in these assays. The HeLa Rad1 with associated nuclease

activity (Rad1 large complex) has not been purified to homogeneity. The co-purification of the nuclease and large Rad1 can be assessed during the course of further chromatographic purification.

Additional experiments involve cleaving hRad1 from GST-hRad1 with Factor Xa and testing purified hRad1 for nuclease activity. Additionally hRad1, FH-hRad1, and GFP-hRad1 are purified from cell extracts and assayed for nuclease activity. With nuclease activity confirmed, hRad1 is mutagenized to inactivate the enzyme. An inactive Rad1 mutant may serve as a dominant negative mutant when expressed in cell lines and enable analysis of cell cycle checkpoint functions of hRad1. An established Rad1 activity assay (exonuclease or DNA-binding, discussed below) is used in a high-throughput screen for hRad1 inhibitors.

Determination of DNA Binding Activity

The DNA-binding properties of GST-Rad1, HeLa Rad1 complexes, and ECR293 epitope-tagged Rad1 (FH-Rad1) were assayed by a nitrocellulose (NC) filter-binding. In this assay, protein and radioactively labeled DNA were mixed and subsequently passed through a nitrocellulose filter which will bind protein but not free DNA. DNA retained on the filter is dependent on the ability of the protein to associate with the DNA. To reduce nonspecific DNA binding, BA85 NC filters (Schleicher & Schuell) were pre-treated in 0.5 M KOH for 10 minutes prior to being equilibrated in 0.1 M Tris-HCl (pH 7.5). Standard DNA binding assays were performed in 20 μl volumes containing 10 mM Tris-HCl (pH 8), 10 mM $MgCl_2$, 0.5 mM DTT, and 1-200 ng ^{32}P-labeled DNA. The DNA was either *Bam*HI-cleaved pUC18 (Random Prime labeled; Boehringer Mannheim) or *Nde*I-cleaved pAdUAS (3′ end labeled as above). DNA was incubated at room temperature or 37°C for 15 minutes in the presence of GST, GST-hRad1-sol6, HeLa Rad1 large complex, HeLa Rad1 small complex, ECR293 FH-Rad1, or appropriate storage buffers (see Examples 3 and 10). Five volumes of 0.1 M Tris-HCl (pH 7.5) was added and the entire sample was passed through a NC filter (drawn by vacuum). The filter was then washed with 0.1 M Tris-HCl (pH 7.5) and allowed to air dry. Counts retained on the filter were measured in a scintillation counter or by use of a PhosphorImager (Molecular Dynamics).

Results indicated that the GST-hRad1-sol6 dialyzed preparation (see

Example 3) had DNA-binding activity which increased linearly with protein concentration (0.1 to 3.5 μM). The substitution of 1 mM EDTA for $MgCl_2$ in the reaction buffer was inhibitory to the DNA-binding activity. In the presence of 0.8 μM protein, DNA-binding increased with increasing DNA concentrations (90 to 900 nM). Heat denaturation of the DNA slightly increased binding, which correlates with the template preference observed in the exonuclease activity assays. GST protein (purified on glutathione beads) also had no significant DNA binding activity.

Results with HeLa Rad1 revealed that both the large and small Rad1 complexes co-purify with a DNA binding activity. The large Rad1 complex had a slight preference for binding heat-denatured DNA and was not sensitive to the replacement of EDTA for $MgCl_2$. The small Rad1 complex preferentially bound double stranded DNA and this activity was inhibited when EDTA replaced magnesium in the incubation buffer. While neither of these Rad1 complexes are purified to homogeneity, these data demonstrate that DNA-binding activity has thus far co-purified with both Rad1 complexes. Demonstration that the DNA-binding activity is inherent to the Rad1 polypeptide or to a Rad1-containing protein complex can be deduced if the activity and Rad1 continue to co-purify chromatographically. Thus far, the elution profile of the small Rad1 complex from Phenyl Superose (Pharmacia Biotech) closely matches the elution profiles of the assayed DNA-binding activities. FH-Rad1 was assayed for binding of denatured DNA and it was found that the peak binding activity corresponded to the peak of FH-Rad1 eluted from the anti-FLAG® mAb column.

Example 9
Rad1 Localization in Meiotic Cells

Indirect immunofluorescence was performed to determine the localization of Rad1 during meiosis in mouse testis. Six monoclonal antibody supernatants (described above) were assayed according to methods previously described [Plug, *et al., Proc. Natl. Acad. Sci.(USA) 94*:1327-1331 (1997)]. Each supernatant recognized mRad1 association with a meiotic structure predicted to be the "dense body" which is a structure that was first observed by electron microscopy [Moses, *Chromosomal 60*:127-137 (1997)] present from mid to late pachytene. The dense body is a structure of unknown function that does not appear to be associated with chromatin but remains at the periphery of the "sex body"

(partially synapsed X and Y chromosomes). The only other protein that has been localized to the dense body is DNA polymerase β [Plug, *et al., supra*] which has DNA repair functions. Rad1 may indeed associate with meiotic chromatin, but the Rad1 monoclonal antibodies may not recognize Rad1 in its chromatin-associated state in these assays.

Example 10
Identification of hRad1 Binding Partners

In order to identify possible binding partner proteins of hRad1, three different protein-protein interaction assays were utilized. In the first assay, the yeast two-hybrid system was used to test for interactions of Rad1 with cloned checkpoint genes (ATR, Chk1, Hus1, and itself) and with mammalian proteins encoded in a cDNA library. In the second assay, GST-Rad1 pull-down assays were performed to detect Rad1-binding proteins present in HeLa and mouse testis extracts. The third strategy was to identify proteins that co-purify with endogenous and epitope-tagged Rad1 from human cell extracts by immuno-detection, mass spectrometry, or peptide sequencing methods.

Rad1 Binding Proteins Identified Using Yeast 2-Hybrid System

The yeast two hybrid system utilized has previously been described [Hollenberg, *et al., Mol. Cell. Biol. 15*:3813-3822 (1995)]. The yeast strain L40, which contains multiple LexA binding sites upstream of HIS3 and beta-galactosidase genes, is transformed with two gene-hybrid vectors that initiate the expression of fusions proteins containing the LexA DNA-binding domain (pBTM116) and the GAL4 transcriptional activation domain (pACT2). Interaction between the LexA fusion protein and the GAL4 fusion protein is scored by expression of HIS3 (growth without histidine) and of beta-galactosidase (blue color in the presence of X-GAL). Additionally, 3-amino triazole can be added to the growth media to make selection for HIS3 expression more stringent. The *Eco*RI/*Not*I hRad1 gene fragment from pGEX-hRad1 was ligated between the *Eco*RI and *Xho*I sites of pBTM116 in the presence of *Xho*I-*Not*I linker oligonucleotides to create pBTM116-Rad1. An *Eco*RI/*Sal*I Chk1 gene fragment, [*Flaggs, et al., Curr.Biol 7*:977-986 (1997)] was ligated into the *Eco*RI/*Sal*I sites of pBTM116 to create pBTM116-Chk1. Additionally, four different gene fragments from ATR were cloned into pBTM116 to create pBTM116-ATR/AB,

-ATR/CD, -ATR/EF, and -ATR/GH. A similar strategy was employed to create pACT2-Rad1 and pACT2-Chk1 (genes were cloned between *Eco*RI/*Xho*I sites of pACT2), except that an EcoRI linker was used to place the hRad1 and Chk1 genes in the correct reading frame. The human Hus1 gene was cloned by PCR amplification from human placental cDNA (Clontech) using primers that hybridized to the ends of the Hus1 coding region and included flanking restriction sites. The primers were HUS1-*Bam*RI (SEQ ID NO: 33) and HUS1-SalPst (SEQ ID NO: 34). The 0.1ml PCR mixture included 100 ng each primer , 0.2 ng cDNA, 1X Perkin-Elmer Taq Polymerase Buffer, 2.5 mM $MgCl_2$, 100 mM dNTP's, and 5 units of AmpliTaq® polymerase (Perkin-Elmer). The PCR mixture was heated 4 min at 94°C, then underwent 30 cycles of 1 min at 94°C, 2 min at 50°C, and 4 min at 72°C. The Hus1 PCR product was gel purified, digested with *Bam*HI and *Sal*I, and ligated between the *Bam*HI and *Xho*I sites of pACT2 to produce pACT2-Hus1. Additionally, two deletion mutations were made from pACT2-Hus1. A vector (pACT2-Hus1-X/Stu) that only encoded the N-terminal 89 amino acids of Hus1 was made by excising the *Xho*I/*Stu*I fragment from pACT2-Hus1, blunting the vector ends, and recircularizing with DNA ligase. A vector (pACT2-Hus1-Sma/X) that only encoded the C-terminal 192 aminoacids of Hus1 was made by excising the *Sma*I/*Xho*I fragment from pACT2-Hus1, blunting the vector ends, and recircularizing with DNA ligase. Correct reading frame of the constructs was verified by DNA sequencing. Expression of the fusion proteins in L40 yeast was verified on Western blots using LexA-specific and Gal4-specific antibodies.

Co-transformation of different plasmid pairs into L40 indicated that hRad1 does not bind to Chk1, the four different ATR domains, or to itself. However, Rad1 did interact with Hus1 in this two hybrid assay. Both the full length hHus1 gene and the truncate expressing the C-terminal 192 amino acids of Hus1 had the ability to bind Rad1. The Hus1 gene truncate encoding the N-terminal 89 amino acids did not interact with Rad1. In *S. pombe* the hus1-Rad1 interaction also has been demonstrated by co-immunoprecipitation experiments [Kostrub, et al, EMBO J. 17:2055-66 (1998)], so it appears that a conserved function of Rad1 is to bind the checkpoint protein Hus1.

Transformation of L40 containing pBTM116-Rad1 with cDNA fusion

libraries (human testis MatchMaker library, Clontech) and mouse embryo cDNA library with the VP16 transcriptional activation domain was performed as previously described to identify hRad1 interacting proteins. Fifty positive interactors were identified from each library which were subjected to further analysis (DNA sequence analysis and specificity tests using the two hybrid system). Three putative Rad1 interacting proteins were identified several times in this screen.

Two of the putative Rad1 interacting proteins identified by this screen were fragments of known genes. The first fragment was identified as mouse BS69 (SEQ ID. NO: 35)(Human Genbank Accession No.: 1362759; Mouse EST Genbank Accession No: AI317203). SEQ ID NO: 36 shows the predicted amino acid sequence of the protein encoded by mouse BS69. The second fragment was identified as mouse BAT1 (SEQ ID. NO.: 37)(Mouse Genbank Accession No: AF118128). SEQ ID NO.: 38 shows the predicted amino acid sequence of the protein encoded by mouse BAT1.

The other putative Rad1 interacting protein identified by this screen was a novel gene referred to as G_4-G_6-H_2-H_3-H_6. Homologous mouse and human genes were also assembled from sequences in EST databases. Aligning mouse ESTs including Accession numbers W65708 (most 5′) and AA710259 (most 3′) resulted in the consensus nucleotide and amino acid sequences set forth in SEQ ID NOs: 39 and 40. Aligning human ESTs including Accession numbers AA359286 (most 5′) and AA470584 (most 3′)resulted in the consensus nucleotide and amino acid sequences set forth in SEQ ID NOs: 41 and 42. SEQ ID NO: 43 is a consensus sequence of G_4-G_6-H_2-H_3-H_6 generated using the five clones containing the G_4-G_6-H_2-H_3-H_6 insert. SEQ ID NO: 44 provides the predicted amino acid sequence of G_4-G_6-H_2-H_3-H_6.

***In vitro* Binding of Mammalian Proteins to GST-hRad1**

In vitro protein binding assays were performed by incubating GST (4.1 μg) or GST-hRad1 (2.3 μg) with HeLa cell extract (400 μg) or mouse testes extract (650 μg) in a volume of 0.1 ml. Binding conditions included 27 μl GST-Rad1 (or GST) diluted in Storage Buffer (see Example 3), 40 μl HeLa cell extract [Finnie, *et al., Proc. Natl. Acad. Sci. (USA) 92*:320-324 (1995)] or 40 μl mouse testes extract [Keegan, *et al., Genes & Dev. 10*:2432-2437 (1996)] and 33 μl buffer A (20 mM Tris HCl, pH 8, 50 mM potassium acetate, 0.05% NP-40, 1 mM Mg acetate, 1 mM DTT,

and 0.2 mM AEBSF). Sequential incubations were at 0°C for 10 minutes, 22°C for 10 minutes, and 0°C for 10 minutes. Samples were gently mixed every 10 minutes. A 50% slurry of glutathione agarose (10μl in Buffer A) was added and incubation continued at 4°C for 30 minutes with occasional mixing. The GSH matrix was sedimented by a five second centrifugation (approximately 15,000 x *g*) and the supernatant was saved in a separate tube. The pellet was washed twice with 200 μl Buffer A. Elution of proteins from the matrix was performed by resuspension in 20 mM glutathione Buffer (see Example 3) and incubation at 22°C for five minutes. Following a five second centrifugation, the eluate was removed to a clean tube. Western analysis of the extracts, supernatants (unbound proteins), and eluates was performed. Proteins were separated on a 8 to 16% polyacrylamide gel (Novex) and transferred to a PVDF membrane. Probing the Western with antibodies was carried out in TBS containing 0.3% Tween 20, and 5% milk. Antibodies to ATR, replication protein A (RPA), c-abl (Ab-2; Calbiochem), and DNA repair proteins MRE11 and Rad50 were used in consecutive probings. Secondary antibodies were either goat anti-mouse IgG or goat anti-rabbit IgG conjugated to horseradish peroxidase (Bio-Rad). Development of Western blots was carried out using ECL reagents (Pierce) followed by autoradiography.

Results indicated that ATR, MRE11 and Rad50 [Delganov, *et al., Mol. Cell. Biol. 16*: 4832-4841 (1996)] may associate with Rad1, as they were slightly more abundant in GST-hRad1-containing eluates than from GST-containing eluates. The same result was observed with HeLa and mouse testis protein-binding reactions. There was no evidence of RPA or c-abl binding to GST-hRad1 (or GST). The Western blots can be probed for other checkpoint proteins, such as Hus1 and Rad9, using appropriate antibodies.

Identification of Proteins that Purify with Rad1

Analysis of proteins that co-purify with endogenous Rad1 (HeLa) and FH-Rad1 (ECR293; see Example 7) was carried out in order to identify proteins that form stable complexes with Rad1 in mammalian cells. The purification of Rad1 (and FH-Rad1) was monitored on Western blots. When purification had reached the stage where Rad1 (or FH-Rad1) could be readily identified in column eluates on silver stained gels, other protein bands that had the same elution profile as Rad1 were

identified and selected for analysis by matrix-assisted laser desorption ionization-mass spectrometry (MALDI-MS), coupled to either HPLC or capillary electrophoresis [Figeys, et al., *Electrophoresis* 19:1811-1818 (1998)]. Larger scale Rad1 and FH-Rad1 preparations can be used to obtain sufficient quantities of co-purifying proteins for identification by peptide sequencing methods (Edman, *Mol Biol Biochem Biophys* 8:211-255 (1970)).

Purification of FH-Rad1 was performed as follows. Stable transfectants expressing FH-Rad1 (see Example 7) were grown to near confluency in T225 flasks and expression of FH-Rad1 was induced by addition of 1mM Muristerone (InVitrogen) for 20-24 hours. Cells were then washed with PBS, scraped from the flasks, centrifuged, and stored as cell pellets at –70°C following rapid freezing in liquid nitrogen. Cell pellets were thawed on ice, then suspended (1ml per T225 flask pellet) in Hypotonic Lysis Buffer (20mM HEPES, pH7.5, 1.5mM $MgCl_2$, 5mM KCl, 1mM DTT, 0.1mM PMSF, 2μg/ml leupeptin) and incubated on ice for 15-30 minutes. The cell suspension was subject to 5 strokes with a tight fitting pestle in a Dounce homogenizer, and incubated 30 min on ice. Cell debris (including nuclei) were pelleted by centrifugation for 10 min at 1700 x *g*. The pellet was designated "nuclear pellet". The supernatant was removed and subjected to a 10 min 20,000 x *g* spin. This supernatant was designated the "cytoplasmic extract". The nuclear pellet was resuspended (0.2ml per T225 flask) in Extraction Buffer (25mM Tris-HCl, pH7.5, 200mM NaCl, 1mM EDTA, 1mM DTT, 10mM Na-bisulfite, 10mM benzamidine, 0.1mM PMSF, 2μg/ml leupeptin) and rocked at 4°C for 30 min. The nuclei suspension was centrifuged 10 min at 20,000 x *g*, and to the saved supernatant 1/100 volume of 5M NaCl was added, and this was designated the "nuclear extract". It was determined that FH-Rad1 (and endogenous Rad1) partitioned nearly equally in the nuclear and cytoplasmic extracts, so each served as sources for FH-Rad1 purification. The cytoplasmic extract was adjusted to 300mM NaCl and 10mM imidazole and batch-loaded onto Ni-NTA (QIAGEN) at 50mg protein per ml matrix, rocking for 1 hour at 4°C. Prior to loading the Ni-NTA was pre-equilibrated in Wash-10 Buffer (20mM HEPES, pH7.5, 300mM NaCl, 10% glycerol, 0.5mM DTT, 0.1mM PMSF, and 10mM imidazole). The matrix/extract slurry was allowed to settle, and the supernatant (unbound protein) was removed. The Ni-NTA matrix (and bound

proteins) was suspended in an equal volume of Wash-10 Buffer and transferred to a Dispo-column (BioRad). The column, which flowed by gravity, was subsequently washed with an additional 3 bed volumes of Wash-10 Buffer and 10 bed volumes of Wash-25 Buffer (same as Wash-10, except with 25mM imidazole). FH-Rad1 was eluted with Wash-250 Buffer (same as Wash-10, except with 250mM imidazole). Peak fractions from the Ni-NTA column were pooled and diluted with an equal volume of Dilution Buffer (50mM Tris-HCl, pH7.5, 10% glycerol, 0.1mM DTT, 0.1mM PMSF). Anti-FLAG® M2 mAb matrix was washed with TBSP (50mM Tris-HCl, pH7.5, 10% glycerol, 150mM NaCl, 0.1mM DTT, 0.1mM PMSF) and added to protein solution (at 1ml per 2mg protein). The M2 matrix rocked with the protein 2 hr at 4°C. The slurry was transferred to a Dispo-column. After the unbound proteins flowed through, the column was washed with 35 bed volumes of TBSP. Elution of FH-Rad1 was performed with 100μg/ml FLAG peptide in TBSP, at a flow rate of 15 min per bed volume. Eluate fractions were stored directly at –70°C following freezing in liquid nitrogen. FH-Rad1 in the nuclear extract was purified over Ni-NTA and anti-FLAG M2 columns by the same protocol, except that rather than adding NaCl prior to binding Ni-NTA, $MgCl_2$ was added to 2mM. Peak FH-Rad1 fractions were analyzed on silver stained gels and FH-Rad1 (~36kDa) was the prominent band in both preparations (nuclear and cytoplasmic). Additionally, co-purifying proteins with the same elution profile as FH-Rad1 were observed. The most prominent of these co-purifying proteins is ~32kDa, and is most likely Hus1 based on its size and the 2 hybrid results (see above). Western blot analysis revealed that endogenous Rad1 does not co-purify with FH-Rad1, which confirms the 2 hybrid result that Rad1 does not bind itself. Several protein bands from the cytoplasmic FH-Rad1 preparation have been excised from silver stained gels and subject to mass spectrometric analysis. Additionally, M2 eluate fractions across the FH-Rad1 peak were tested for DNA-binding activity (see Example 8).

Purification of endogenous Rad1 from HeLa cells revealed that there are at least two distinct forms of Rad1 (designated "small' and "large" Rad1 complexes) that can be chromatographically separated. Purification of these Rad1 complexes proceeded as follows. HeLa cell pellets (Cellex Biosciences) were thawed on ice and suspended in (4.5ml per cell gram) Hypotonic Lysis Buffer (as above, with

2μg/ml aprotinin). After 15 min on ice the suspension was subject to 20 strokes with a tight fitting pestle in a Dounce homogenizer, and incubated 30-60 min longer on ice. The suspension was the centrifuged at 1700 x g for 10 min, and the supernatant was centrifuged again at 12,000 x g for 10 min to prepare the "cytoplasmic extract". The cytoplasmic extract was centrifuged at 100,000 x g for 1 hour, and the supernatant (S100) was saved. The S100 was subject to sequential ammonium sulfate (AS) precipitations at 20%, 40%, and 70% saturating AS levels. Following an hour of stirring the S100/AS solution on ice, the mixture was centrifuged at 12,000 x g for 20 minutes. The 40% and 70% AS precipitates contained the majority of the Rad1. The 40% AS precipitate was resuspended (0.25 of S100 volume) in Buffer A-25 (25mM Tris-HCl, pH7.5, 10% glycerol, 0.01% NP-40, 25mM NaCl, 1mM DTT, 10mM Na-bisulfite. 0.5μg/ml leupeptin, 11mM benzamidine, 0.1mM PMSF, 1mM EDTA) and dialyzed into Buffer A-200 (above buffer with 200mM NaCl). The protein was then loaded on a Heparin Sepharose FF column (Pharmacia Biotech) at about 20mg protein per ml matrix. Proteins were eluted off with a 0.2-1M NaCl gradient in Buffer A (without EDTA). About half the Rad1 flowed through the column and the remainder eluted at about 400mM NaCl. The flow through fractions were stored at –70°C, awaiting further analysis. The eluted Rad1 peak fractions were pooled and dialyzed into Q Buffer (25mM Tris-HCl, pH8, 0.1mM EDTA, 10% glycerol, 1mM DTT, 0.1 mM PMSF, 0.01% NP-40) containing 100mM NaCl, 0.1mM EGTA, and 0.5mg/ml leupeptin. The dialyzed protein was loaded on a Mono Q column (Pharmacia Biotech) and eluted with a 100-600mM NaCl gradient in Q Buffer. Rad1 eluted from Mono Q in two distinct peaks, at about 210mM and 260mM NaCl, respectively. Approximately 0.9ml of each Mono Q Rad1 peak was subject to gel filtration chromatography using Superdex 200 (Pharmacia Biotech) that was previously calibrated with molecular weight standards and equilibrated in SD200 Buffer (25mM Tris-HCl, pH8, 0.1mM EDTA, 10% glycerol, 1mM DTT, 0.1mM PMSF, 0.01% NP-40, 200mM NaCl, and 0.5μg/ml leupeptin). The elution profiles of Rad1 from Superdex 200 revealed that the two different Mono Q peaks of Rad1 represent different molecular weight Rad1 complexes. The Rad1 complex that elutes at ~210mM NaCl from Mono Q elutes as a ~40-48kDa protein off Superdex 200, and has been designated the "Rad1 small complex". The Rad1 complex that elutes at

~260mM NaCl from Mono Q elutes as a ~76-86kDa protein off Superdex 200, and has been designated the "Rad1 large complex". Because protein shape also affects protein retention of these columns, gel filtration does not always accurately predict the molecular weight of proteins. However, from these results it can be predicted that the Rad1 large complex is not a Rad1 monomer, and probably contains Rad1 in association with at least one other protein. Further purification of the Rad1 large complex can reveal the protein(s) that is associated with Rad1 in this complex. The Rad1 small complex has been further purified by pooling the Mono Q peak (also done with the Superdex 200 peak) and precipitating with 33%AS to remove the majority of the contaminating proteins. The Rad1-containing 33%AS supernatant was loaded on a Phenyl Superose column (Pharmacia Biotech) and eluted with a gradient of 1.4M ammonium sulfate (with 42mM sodium phosphate, pH 7.5) to water. Rad1 small complex eluted in a single peak at ~130-20mM AS. The peak fractions were adjusted to 1mM DTT and stored at –70°C. Silver stain analysis of the Phenyl Superose fractions of the Rad1 small complex reveal that the Rad1 protein band can be readily detected in the peak fractions, as well as several smaller proteins (most prominently a ~23kDa protein) with the same elution profile. These smaller proteins have been subject to analysis by mass spectrometry. It is possible that Hus1 is a constituent of the Rad1 small complex and that it co-migrates with the Rad1 band (each are ~32kDa), making it undetectable upon silver staining of the 1D protein gels. Therefore, the 32kDa band has also been subjected to analysis by mass spectrometry. The Rad1 large complex can be purified over Phenyl Superose following a 33%AS cut.

What is claimed is:

1. A purified and isolated human Rad1 polypeptide.

2 The polypeptide according to claim 1 comprising the Rad1 amino acid sequence set out in SEQ ID NO: 2.

3. A polynucleotide encoding the polypeptide according to claim 1 or 2.

4. The polynucleotide according to claim 3 comprising the sequence set forth in SEQ ID NO: 1.

5. A polynucleotide encoding a human Rad1 polypeptide selected from the group consisting of:

a) the polynucleotide according to claim 2;

b) a DNA which hybridizes under highly stringent conditions to the complement of the polynucleotide of (a), and

c) a DNA which would hybridize to the polynucleotide of (a) but for degeneracy of the genetic code.

6. The Rad1 polypeptide encoded by the polynucleotide of claim 5.

7. The polynucleotide of claim 5 which is a DNA molecule.

8. The DNA of claim 7 which is a cDNA molecule.

9. The DNA of claim 7 which is a wholly or partially chemically synthesized DNA molecule.

10. An anti-sense polynucleotide which specifically hybridizes with

the complement of the polynucleotide of claim 5.

11. A expression construct comprising the polynucleotide according to claim 5.

12. A host cell transformed or transfected with the polynucleotide according to claim 11.

13. A method for producing a human Rad1 polypeptide comprising the steps of:

a) growing the host cell according to claim 12 under conditions appropriate for expression of the human Rad1 polypeptide and
b) isolating the Rad1 polypeptide from the host cell or the medium of its growth.

14. An antibody specifically immunoreactive with the polypeptide according to claim 1 or 2.

15. The antibody according to claim 14 which is a monoclonal antibody.

16. An immortal cell line which secretes the antibody according to claim 15.

17. An anti-idiotype antibody specifically immunoreactive with the antibody according to claim 14.

18. An antibody according to claim 14 secreted by the hybridoma cell line 267B or 267H.

19. A method to identify a specific binding partner compound of the human Rad1 polypeptide according to claim 1 or 2 comprising the steps of:

a) contacting the Rad1 polypeptide with a compound under conditions which permit binding between the compound and the Rad1 polypeptide;

b) detecting binding of the compound to the Rad1 polypeptide; and

c) identifying the compound as a specific binding partner of the Rad1 polypeptide.

20. The method according to claim 19 wherein the specific binding partner modulates activity of the Rad1 polypeptide.

21. The method according to claim 20 wherein the compound inhibits activity of the Rad1 polypeptide.

22. The method according to claim 20 wherein the compound enhances activity of the Rad1 polypeptide.

23. A compound identified by the method according to claim 19.

24. A method to identify a specific binding partner compound of the Rad1 polynucleotide according to claim 5 comprising the steps of:

a) contacting the Rad1 polynucleotide with a compound under conditions which permit binding between the compound and the Rad1 polynucleotide;

b) detecting binding of the compound to the Rad1 polynucleotide; and

c) identifying the compound as a specific binding partner of the Rad1 polynucleotide.

25. The method according to claim 24 wherein the specific binding partner modulates expression of a Rad1 polypeptide encoded by the Rad1 polynucleotide.

26. The method according to claim 25 wherein the compound inhibits expression of the Rad1 polypeptide.

27. The method according to claim 25 wherein the compound enhances expression of the Rad1 polypeptide.

28. A compound identified by the method according to claim 24.

29. A composition comprising the compound according to claim 28 and a pharmaceutically acceptable carrier

30. A method to identify modulators of Rad1 binding to a ligand comprising the steps of:

a) contacting a Rad1 polypeptide and a ligand in the presence and absence of a compound;

b) detecting Rad1 polypeptide binding to the ligand in the presence and absence of the compound; and

c) identifying the compound as a modulator of Rad1 polypeptide binding to the ligand by detecting a change in Rad1 binding to the polypeptide in the presence of the compound.

31. The method according to claim 29 where the compound is identified as an inhibitor of Rad1 binding to the ligand when decreased Rad1 polypeptide binding to the ligand is detected in the presence of the compound.

32. The method according to claim 30 wherein said ligand is selected from the group consisting of Hus1, BS69, BAT1, MRE11 and Rad50.

33. The method according to claim 30 where the compound is identified as an activator of Rad1 binding to the ligand when increased Rad1 polypeptide binding to the ligand is detected in the presence of the compound.

34. A compound identified by the method according to any one of claims 30, 31, and 33.

35. A composition comprising the compound according to claim 34 and a pharmaceutically acceptable carrier.

36. A purified and isolated binding partner of human Rad1 polypeptide comprising the sequence set forth in SEQ ID NO: 44.

37. A polynucleotide encoding the polypeptide according to claim 36.

38. A polynucleotide according to claim 37 comprising the sequence set forth in SEQ ID NO: 43.

39. A polynucleotide encoding a binding partner of human Rad1 polypeptide selected from the group consisting of:

(a) the polynucleotide according to claim 37;

(b) a DNA which hybridizes under highly stringent conditions to the complement of the polynucleotide of (a);

(c) a DNA which would hybridize to the polynucleotide of (a) but for degeneracy of the genetic code.

40. The Rad1 polypeptide encoded by the polynucleotide of claim 39.

41. The polynucleotide of claim 39 which is a DNA molecule.

42. The DNA of claim 41 which is a cDNA molecule.

43. The DNA of claim 41 which is a wholly or partially chemically synthesized DNA molecule.

44. An anti-sense polynucleotide which specifically hybridizes with the complement of the polynucleotide of claim 40.

45. A expression construct comprising the polynucleotide according to claim 40.

46. A host cell transformed with the polynucleotide according to claim 45.

47. An antibody specifically immunoreactive with the polypeptide according to claim 36 or 37.

48. The antibody according to claim 47 which is a monoclonal antibody.

49. A hybridoma which secretes the antibody according to claim 47.

50. An anti-idiotype antibody specifically immunoreactive with the antibody according to claim 47.

SEQUENCE LISTING

<110> ICOS CORPORATION (Applicant)
Herendeen, Daniel R.(Inventor)

<120> HUMAN RAD1 CELL CYCLE CHECKPOINT

<130> 27866/35406 PCT

<140>
<141>

<150> 09/049,792
<151> 1998-03-27

<160> 44

<170> PatentIn Ver. 2.0

<210> 1
<211> 846
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (1)..(846)

<400> 1

```
atg ccc ctt ctg acc caa cag atc caa gac gag gat gat cag tac agc   48
Met Pro Leu Leu Thr Gln Gln Ile Gln Asp Glu Asp Asp Gln Tyr Ser
  1               5                  10                  15

ctt gtg gcc agc ctt gac aac gtt agg aat ctc tcc act atc ttg aaa   96
Leu Val Ala Ser Leu Asp Asn Val Arg Asn Leu Ser Thr Ile Leu Lys
             20                  25                  30

gct att cat ttc cga gaa cat gcc acg tgt ttc gca act aaa aat ggt   144
Ala Ile His Phe Arg Glu His Ala Thr Cys Phe Ala Thr Lys Asn Gly
         35                  40                  45

atc aaa gta aca gtg gaa aat gca aag tgt gtg caa gca aat gct ttt   192
Ile Lys Val Thr Val Glu Asn Ala Lys Cys Val Gln Ala Asn Ala Phe
     50                  55                  60

att cag gct gga ata ttt cag gag ttt aaa gtt cag gaa gag tct gtt   240
Ile Gln Ala Gly Ile Phe Gln Glu Phe Lys Val Gln Glu Glu Ser Val
 65                  70                  75                  80

act ttt cga att aat tta act gtc ctt tta gac tgt tta tct att ttt   288
Thr Phe Arg Ile Asn Leu Thr Val Leu Leu Asp Cys Leu Ser Ile Phe
                 85                  90                  95
```

```
gga tca agt cct atg cca ggg act tta act gca ctt cga atg tgt tac   336
Gly Ser Ser Pro Met Pro Gly Thr Leu Thr Ala Leu Arg Met Cys Tyr
            100                 105                 110

caa ggt tat ggt tac cct ttg atg ctg ttc ctg gaa gaa gga gga gtg   384
Gln Gly Tyr Gly Tyr Pro Leu Met Leu Phe Leu Glu Glu Gly Gly Val
        115                 120                 125

gtg aca gtc tgc aaa atc aat aca cag gaa cct gag gag acc ctg gac   432
Val Thr Val Cys Lys Ile Asn Thr Gln Glu Pro Glu Glu Thr Leu Asp
    130                 135                 140

ttt gat ttc tgc agc acc aat gtt att aat aaa att att ctg cag tca   480
Phe Asp Phe Cys Ser Thr Asn Val Ile Asn Lys Ile Ile Leu Gln Ser
145                 150                 155                 160

gag ggg ctc cgt gaa gca ttt tct gaa ttg gat atg acg agt gaa gtc   528
Glu Gly Leu Arg Glu Ala Phe Ser Glu Leu Asp Met Thr Ser Glu Val
                165                 170                 175

cta caa att acc atg tct cct gac aag cct tat ttc agg tta tct act   576
Leu Gln Ile Thr Met Ser Pro Asp Lys Pro Tyr Phe Arg Leu Ser Thr
            180                 185                 190

ttt gga aat gca gga agt tcc cac ctt gac tat ccc aaa gat tct gat   624
Phe Gly Asn Ala Gly Ser Ser His Leu Asp Tyr Pro Lys Asp Ser Asp
        195                 200                 205

ttg atg gaa gca ttt cat tgt aat cag acc caa gtc aac aga tac aag   672
Leu Met Glu Ala Phe His Cys Asn Gln Thr Gln Val Asn Arg Tyr Lys
    210                 215                 220

att tcc tta ctg aaa ccc tct aca aag gca tta gtc cta tct tgt aag   720
Ile Ser Leu Leu Lys Pro Ser Thr Lys Ala Leu Val Leu Ser Cys Lys
225                 230                 235                 240

gta tct att cgg aca gat aac aga ggc ttc ctt tca tta cag tat atg   768
Val Ser Ile Arg Thr Asp Asn Arg Gly Phe Leu Ser Leu Gln Tyr Met
                245                 250                 255

att aga aat gaa gat gga caa ata tgt ttt gtg gaa tat tac tgc tgc   816
Ile Arg Asn Glu Asp Gly Gln Ile Cys Phe Val Glu Tyr Tyr Cys Cys
            260                 265                 270

cct gat gaa gaa gtt cct gaa tct gag tct                           846
Pro Asp Glu Glu Val Pro Glu Ser Glu Ser
        275                 280
```

<210> 2
<211> 282
<212> PRT
<213> Homo sapiens

<400> 2

```
Met Pro Leu Leu Thr Gln Gln Ile Gln Asp Glu Asp Asp Gln Tyr Ser
  1               5                  10                  15

Leu Val Ala Ser Leu Asp Asn Val Arg Asn Leu Ser Thr Ile Leu Lys
            20                  25                  30

Ala Ile His Phe Arg Glu His Ala Thr Cys Phe Ala Thr Lys Asn Gly
        35                  40                  45

Ile Lys Val Thr Val Glu Asn Ala Lys Cys Val Gln Ala Asn Ala Phe
    50                  55                  60

Ile Gln Ala Gly Ile Phe Gln Glu Phe Lys Val Gln Glu Glu Ser Val
65                  70                  75                  80

Thr Phe Arg Ile Asn Leu Thr Val Leu Leu Asp Cys Leu Ser Ile Phe
                85                  90                  95

Gly Ser Ser Pro Met Pro Gly Thr Leu Thr Ala Leu Arg Met Cys Tyr
            100                 105                 110

Gln Gly Tyr Gly Tyr Pro Leu Met Leu Phe Leu Glu Glu Gly Gly Val
        115                 120                 125

Val Thr Val Cys Lys Ile Asn Thr Gln Glu Pro Glu Glu Thr Leu Asp
    130                 135                 140

Phe Asp Phe Cys Ser Thr Asn Val Ile Asn Lys Ile Ile Leu Gln Ser
145                 150                 155                 160

Glu Gly Leu Arg Glu Ala Phe Ser Glu Leu Asp Met Thr Ser Glu Val
                165                 170                 175

Leu Gln Ile Thr Met Ser Pro Asp Lys Pro Tyr Phe Arg Leu Ser Thr
            180                 185                 190

Phe Gly Asn Ala Gly Ser Ser His Leu Asp Tyr Pro Lys Asp Ser Asp
        195                 200                 205

Leu Met Glu Ala Phe His Cys Asn Gln Thr Gln Val Asn Arg Tyr Lys
    210                 215                 220

Ile Ser Leu Leu Lys Pro Ser Thr Lys Ala Leu Val Leu Ser Cys Lys
225                 230                 235                 240

Val Ser Ile Arg Thr Asp Asn Arg Gly Phe Leu Ser Leu Gln Tyr Met
                245                 250                 255

Ile Arg Asn Glu Asp Gly Gln Ile Cys Phe Val Glu Tyr Tyr Cys Cys
            260                 265                 270

Pro Asp Glu Glu Val Pro Glu Ser Glu Ser
        275                 280
```

<210> 3
<211> 840
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (1)..(840)

<400> 3

```
atg cct ctc cta acc cag tac aat gaa gag gag tac gaa cag tac tgc   48
Met Pro Leu Leu Thr Gln Tyr Asn Glu Glu Glu Tyr Glu Gln Tyr Cys
  1               5                  10                  15

tta gtg gcc agc ctt gac aac gtt agg aat ctc ttc act gtc ttg aaa   96
Leu Val Ala Ser Leu Asp Asn Val Arg Asn Leu Phe Thr Val Leu Lys
            20                  25                  30

gcc att cat ttc aga gaa cac gcc acg tgt ttt gct acc aaa aac gga   144
Ala Ile His Phe Arg Glu His Ala Thr Cys Phe Ala Thr Lys Asn Gly
         35                  40                  45

atc aag gtt aca gtg gag aat gca aag tgt gtg caa gca aat gcc ttt   192
Ile Lys Val Thr Val Glu Asn Ala Lys Cys Val Gln Ala Asn Ala Phe
     50                  55                  60

att cag gct gac gtg ttt cag gaa ttt gtc att cag gaa gaa tct gtt   240
Ile Gln Ala Asp Val Phe Gln Glu Phe Val Ile Gln Glu Glu Ser Val
 65                  70                  75                  80

act ttt cga att aac tta act atc ctt tta gac tgt tta tct att ttt   288
Thr Phe Arg Ile Asn Leu Thr Ile Leu Leu Asp Cys Leu Ser Ile Phe
                85                  90                  95

gga tca agt cct aca cca ggg act ttg act gcg ctt cgg atg tgt tac   336
Gly Ser Ser Pro Thr Pro Gly Thr Leu Thr Ala Leu Arg Met Cys Tyr
            100                 105                 110

caa ggt tat ggt cac cca ctg atg cta ttt cta gaa gaa gga gga gtg   384
Gln Gly Tyr Gly His Pro Leu Met Leu Phe Leu Glu Glu Gly Gly Val
        115                 120                 125

gtg acg gtc tgc aaa att acc act cag gag cct gag gag aca ctg gat   432
Val Thr Val Cys Lys Ile Thr Thr Gln Glu Pro Glu Glu Thr Leu Asp
    130                 135                 140

ttt gat ttc tgc agc acc aat gtt atg aat aaa att atc ctg cag tca   480
Phe Asp Phe Cys Ser Thr Asn Val Met Asn Lys Ile Ile Leu Gln Ser
145                 150                 155                 160

gag ggg ctc cgg gaa gcc ttt tct gag ctg gac atg aca ggt gat gtc   528
Glu Gly Leu Arg Glu Ala Phe Ser Glu Leu Asp Met Thr Gly Asp Val
                165                 170                 175
```

```
cta cag atc act gtg tct cct gac aag ccc tat ttc agg ttg tct act   576
Leu Gln Ile Thr Val Ser Pro Asp Lys Pro Tyr Phe Arg Leu Ser Thr
            180                 185                 190

ttt gga aat gca gga aac tcc cat ctt gac tat ccc aaa gat tcc gac   624
Phe Gly Asn Ala Gly Asn Ser His Leu Asp Tyr Pro Lys Asp Ser Asp
        195                 200                 205

ttg gtg gaa gcc ttt cac tgt gat aag acc cag gtc aac aga tac aag   672
Leu Val Glu Ala Phe His Cys Asp Lys Thr Gln Val Asn Arg Tyr Lys
    210                 215                 220

ctg tcg cta ctg aag ccc tct aca aag gca cta gct tta tcc tgt aaa   720
Leu Ser Leu Leu Lys Pro Ser Thr Lys Ala Leu Ala Leu Ser Cys Lys
225                 230                 235                 240

gtg tct atc cgg aca gat aac cga ggc ttc ctc tcc tta cag tac atg   768
Val Ser Ile Arg Thr Asp Asn Arg Gly Phe Leu Ser Leu Gln Tyr Met
                245                 250                 255

att aga aat gaa gat ggg cag ata tgt ttt gtg gaa tat tac tgc tgc   816
Ile Arg Asn Glu Asp Gly Gln Ile Cys Phe Val Glu Tyr Tyr Cys Cys
            260                 265                 270

cct gat gaa gaa gtt cct gag tct                                  840
Pro Asp Glu Glu Val Pro Glu Ser
        275                 280
```

<210> 4
<211> 280
<212> PRT
<213> Mus musculus

<400> 4

```
Met Pro Leu Leu Thr Gln Tyr Asn Glu Glu Glu Tyr Glu Gln Tyr Cys
  1               5                  10                  15

Leu Val Ala Ser Leu Asp Asn Val Arg Asn Leu Phe Thr Val Leu Lys
             20                  25                  30

Ala Ile His Phe Arg Glu His Ala Thr Cys Phe Ala Thr Lys Asn Gly
         35                  40                  45

Ile Lys Val Thr Val Glu Asn Ala Lys Cys Val Gln Ala Asn Ala Phe
     50                  55                  60

Ile Gln Ala Asp Val Phe Gln Glu Phe Val Ile Gln Glu Glu Ser Val
 65                  70                  75                  80

Thr Phe Arg Ile Asn Leu Thr Ile Leu Leu Asp Cys Leu Ser Ile Phe
                 85                  90                  95

Gly Ser Ser Pro Thr Pro Gly Thr Leu Thr Ala Leu Arg Met Cys Tyr
            100                 105                 110
```

```
Gln Gly Tyr Gly His Pro Leu Met Leu Phe Leu Glu Glu Gly Gly Val
        115                 120                 125

Val Thr Val Cys Lys Ile Thr Thr Gln Glu Pro Glu Glu Thr Leu Asp
    130                 135                 140

Phe Asp Phe Cys Ser Thr Asn Val Met Asn Lys Ile Ile Leu Gln Ser
145                 150                 155                 160

Glu Gly Leu Arg Glu Ala Phe Ser Glu Leu Asp Met Thr Gly Asp Val
                165                 170                 175

Leu Gln Ile Thr Val Ser Pro Asp Lys Pro Tyr Phe Arg Leu Ser Thr
            180                 185                 190

Phe Gly Asn Ala Gly Asn Ser His Leu Asp Tyr Pro Lys Asp Ser Asp
        195                 200                 205

Leu Val Glu Ala Phe His Cys Asp Lys Thr Gln Val Asn Arg Tyr Lys
    210                 215                 220

Leu Ser Leu Leu Lys Pro Ser Thr Lys Ala Leu Ala Leu Ser Cys Lys
225                 230                 235                 240

Val Ser Ile Arg Thr Asp Asn Arg Gly Phe Leu Ser Leu Gln Tyr Met
                245                 250                 255

Ile Arg Asn Glu Asp Gly Gln Ile Cys Phe Val Glu Tyr Tyr Cys Cys
            260                 265                 270

Pro Asp Glu Glu Val Pro Glu Ser
        275                 280
```

<210> 5
<211> 551
<212> DNA
<213> Homo sapiens

<400> 5
gtggagggcc ggtctgaaga gtggcgggac tggcttcact tcctccgcgg ttcannggag 60

ccgcctcgct cctcttcagg gactttgctg agaagggctc tcgggcgtcc agaccccacc 120

gcaaaggtgt ttggcgatcc gcgagaagtt gttggcccca ggagcatncc tcggggcnga 180

atgcgcagtg gacgatgccc cttctgaccc aacagatcca agacgaggat gatcagtaca 240

gccttgtggc cagccttgac aacgttagga atctctccac tatcttgaaa gctattcatt 300

tccgagaaca atgccacgtg tttcgcaact aaaaatggta tcaaagtaac agttggaaaa 360

tgcaaagtgt gtgcaagcaa atgcttttat tcagggactt taacttgcac ttcgaatgtg 420

ttaccaaggt tatggtttac ccttttgatg ctgttcctgg aaagaaggan gagtggtgac 480

agtctgcaaa ntcaatacac aggaacctga ggagacctgg actttgattt ctgcagcacc 540

aatgttatta a 551

<210> 6
<211> 484
<212> DNA
<213> Homo sapiens

<400> 6
atgactacag gaaaacattt atttacatgc cctctacaaa atggatttac aaaacatagt 60

aactattagg gtacatgacc ttgctcctat cttccccatt gtgcttcttc tctatagaaa 120

atccaatatg aaatgacaaa gagtactgta ctcagaataa gaacttcatc tatcataaat 180

gtacacataa atatcagtga attgtcatac tcaagactca gattcaggaa cttcttcatc 240

agggcagcag taatattcca caaaacatat ttgtccatct tcatttctaa tcatatactg 300

taatgaaagg aagcctctgt tatctgtccg aatagatacc ttacaagata ggactaatgc 360

ctttgtagag ggtttcagta aggaaatctt gtatctgttg acttgggtct gattacaatg 420

aaatgcttcc atcaaatcag aatctttggg atagtcaagg tgggaacttc ctgcatttcc 480

aaaa 484

<210> 7
<211> 362
<212> DNA
<213> Homo sapiens

<400> 7
aggcttcctt tcattacagt atatgattag aaatgaagat ggacaaatat gttttgtgga 60

atattactgc tgccctgatg aagaagttcc tgaatctgag tcttgagtat gacaattcac 120

tgatatttat gtgtacattt atgatagatg aagttcttat tctgagtaca gtactctttg 180

tcatttcata ttgggatttt ctatagagaa ggaagcacaa tgggggggaag gatagggagc 240

aagggtcatg tacccttant agttacttat gttttgttaa aatccctttt gtaggagggg 300

ccatgttaaa taantgtttt ccctgttagg tcctaggntt attccggggc tgtcccntta 360

gg 362

<210> 8
<211> 350
<212> DNA
<213> Homo sapiens

<400> 8
taattgccta ggctggagtg caatggcgtc atctcagctc accacaacct ccgcctcccg 60

ggttcaagca attctcctgc ctcagcctcc cgagtagctg gcattacagg catgtgccac 120

cacgtccggc taatttttg tatttttagt agagatgggg tttctccacg ttggncaggc 180

tggtctcgaa ctcctgacct caggtgancc gcccacctcg gcttcccaaa gtgctgggat 240

taaagccacc acacccggcc acctttagac tcttgaaatg tnacgctgac tcacaattat 300

tcccatgagt tcaaaagaca gaactaaagg acagtcctga ataatctatg 350

<210> 9
<211> 423
<212> DNA
<213> Homo sapiens

<400> 9
agtggcggga ctggcttcac ttcctccgcg gttcctcgga gccgcctcgc tcctcttcag 60

ggactttgct gagaagggct ctcgggcgtc cagaccccac cgcaaaggtg tttggcgatc 120

cgcgagaagt tgttggcccc aggagcatcc ctcggggccg aatgcgcagt ggacgatgcc 180

ccttctgacc caacagatcc aagacgagga tgatcagtac agccttgtgg cagccttgac 240

aacgttagga atctctccac tatcttgaaa gctattcatt tccgagaaca tgccacgtgt 300

ttcgcaacta aaaatggtat caaagtaaca gtggaaaatg caaagtgtgt gcaagcaaat 360

gcttttattc aggctggaat atttcaggag tttaaagttc aggaagagtc tgttactttt 420

cga 423

<210> 10
<211> 433
<212> DNA
<213> Homo sapiens

<400> 10
atctatgact acaggaaaac atttatttac atgccctcta caaaatggat ttacaaaaca 60

tagtaactat tagggtacat gaccttgctc ctatcttccc cattgtgctt cttctctata 120

gaaaatccaa tatgaaatga caaagagtac tgtactcaga ataagaactt catctatcat 180

aaatgtacac ataaatatca gtgaattgtc atactcaaga ctcagattca ggaacttctt 240

catcagggca gcagtaatat tccacaaaac atatttgtcc atcttcattt ctaatcatat 300

actgtaatga aaggaagcct ctgttatctg tccgaataga taccttacaa gataggacta 360

atgcctttgt agagggtttc agtaagggaa tcttgtatct gttgacttgg gtctgattac 420

aatgaaatgc ttc 433

<210> 11
<211> 503
<212> DNA
<213> Mus musculus

<400> 11
tggaccatcc atgcctctcc taacccagta caatgaagag gagtacgaac agtactgctt 60

agtggccagc cttgacaacg ttaggaatct cttcactgtc ttgaaagcca ttcatttcag 120

agaacacgcc acgtgttttg ctaccaaaaa cggaatcaag gttacagtgg agaatgcaaa 180

gtgtgtgcaa gcaaatgcct ttattcaggc tgacgtgttt caggaatttg tcattcagga 240

agaatctgtt acttttcgaa ttaacttaac tatcctttta gactgtttat ctatttttgg 300

atcaagtcct acaccaggga ctttgactgc gcttcggatg tgttaccaag gttatggtca 360

cccactgatg ctatttctag aagaaggagg agtggtgacg gtctgcaaaa ttaccactca 420

ggagcctgag gagacactgg attttgattt ctgcagcacc aatgttatga ataaaattat 480

cctgcagtca gaggggctcc gga 503

<210> 12
<211> 499
<212> DNA
<213> Mus musculus

<400> 12
tggaccatcc atgcctctcc taacccagta caatgaagag gagtacgaac agtactgctt 60

agtggccagc cttgacaacg ttaggaatct cttcactgtc ttgaaagcca ttcatttcag 120

agaacacgcc acgtgttttg ctaccaaaaa cggaatcaag gttacagtgg agaatgcaaa 180

gtgtgtgcaa gcaaatgcct ttattcaggc tgacgtgttt caggaatttg tcattcagga 240

agaatctgtt acttttcgaa ttaacttaac tatcctttta gactgtttat ctatttttgg 300

atcaagtcct acaccaggga ctttgactgc gcttcggatg tgttaccaag gttatggtca 360

cccactgatg ctatttctag aagaaggagg agtggtgacg gtctgcaaaa ttaccactca 420

ggagcctgag gagacactgg attttgattt ctgcagcacc aatgttatga ataaaattat 480

cctgcagtca gaggggctc 499

<210> 13
<211> 418
<212> DNA
<213> Mus musculus

<400> 13
tggaccatcc atgcctctcc taacccagta caatgaagag gagtacgaac agtactgctt 60

agtggccagc cttgacaacg ttaggaatct cttcactgtc ttgaaagcca ttcatttcag 120

agaacacgcc acgtgttttg ctaccaaaaa cggaatcaag gttacagtgg agaatgcaaa 180

gtgtgtgcaa gcaaatgcct ttattcaggc tgacgtgttt caggaatttg tcattcagga 240

agaatctgtt acttttcgaa ttaacttaac tatcctttta gactgtttat ctatttttgg 300

atcaagtcct acaccaggga ctttgactgc gcttcggatg tgttaccaag gttatggtca 360

cccactgatg ctatttctag aagaaggagg agtggtgacg gtctgcaaaa ttaccact 418

<210> 14
<211> 1145
<212> DNA
<213> Mus musculus

<400> 14
gtggagggcc ggtctgaaga gtggcgggac tggcttcact tcctccgcgg ttcctcggag 60

ccgcctcgct cctcttcagg gactttgctg agaagggctc tcgggcgtcc agaccccacc 120

gcaaaggtgt ttggcgatcc gccgagaagt tgttggcccc aggagcatcc ctcggggccg 180

aatgcgcagt ggacgatgcc ccttctgacc caacagatcc aagacgagga tgatcagtac 240

agccttgtgg ccagccttga caacgttagg aatctctcca ctatcttgaa agctattcat 300

ttccgagaac atgccacgtg tttcgcaact aaaaatggta tcaaagtaac agtggaaaat 360

gcaaagtgtg tgcaagcaaa tgcttttatt cagggacttt aactgcactt cgaatgtgtt 420

accaaggtta tggttaccct ttgatgctgt tcctggaaga aggaggagtg gtgacagtct 480

gcaaaatcaa tacacaggaa cctgaggaga ccctggactt tgatttctgc agcaccaatg 540

ttattaataa aattattctg cagtcagagg ggctccgtga agcattttct gaattggata 600

tgacgagtga agtcctacaa attaccatgt ctcctgacaa gccttatttc aggttatcta 660

cttttggaaa tgcaggaagt tcccaccttg actatcccaa agattctgat ttgatggaag 720

catttcattg taatcagacc caagtcaaca gatacaagat ttccttactg aaaccctcta 780

caaaggcatt agtcctatct tgtaaggtat ctattcggac agataacaga ggcttccttt 840

cattacagta tatgattaga aatgaagatg gacaaatatg ttttgtggaa tattactgct 900

gccctgatga agaagttcct gaatctgagt cttgagtatg acaattcact gatatttatg 960

tgtacattta tgatagatga agttcttatt ctgagtacag tactctttgt catttcatat 1020

tggattttct atagagaaga agcacaatgg ggaagatagg agcaaggtca tgtaccctaa 1080

tagttactat gttttgtaaa tccattttgt agagggcatg taaataaatg ttttcctgta 1140

gtcat 1145

<210> 15
<211> 787
<212> DNA
<213> Mus musculus

<400> 15
tttcattgta atcagaccca agtcaacaga tacaagattt ccttactgaa accctctaca 60

aaggcattag tcctatcttg taaggtatct attcggacag ataacagagg cttcctttca 120

ttacagtata tgattagaaa tgaagatgga caaatatgtt ttgtggaata ttactgctgc 180

cctgatgaag aagttcctga atctgagtct tgagtatgac aattcactga tatttatgtg 240

tacatttatg atagatgaag ttcttattct gagtacagta ctctttgtca tttcatattg 300

gattttctat agagaagaag cacaatgggg aagataggag caaggtcatg taccctaata 360

gttactatgt tttgtaaatc cattttgtag agggcatgta aataaatgtt ttcctgtagt 420

catagattat tcaggactgt cctttagttc tgtcttttga actcatggga ataattgtga 480

gtcagcgtaa catttcaaga gtctaaaggt ggccgggtgt ggtggcttta atcccagcac 540

tttgggaagc cgaggtgggc ggatcacctg aggtcaggag ttcgagacca gcctgaccaa 600

cgtggagaaa ccccatctct actaaaaata caaaaaatta gccggacgtg gtggcacatg 660

cctgtaatgc cagctactcg ggaggctgag gcaggagaat tgcttgaacc cgggaggcgg 720

aggttgtggt gagctgagat gacgccattg cactccagcc taggcaataa gagcaaaact 780

ctgcctc 787

<210> 16
<211> 1240
<212> DNA
<213> Mus musculus

<400> 16
agtggcggga ctggcttcac ttcctccgcg gttcctcgga gccgcctcgc tcctcttcag 60

ggactttgct gagaagggct ctcgggcgtc cagaccccac cgcaaaggtg tttggcgatc 120

cgccgagaag ttgttggccc caggagcatc cctcggggcc gaatgcgcag tggacgatgc 180

cccttctgac ccaacagatc caagacgagg atgatcagta cagccttgtg gccagccttg 240

acaacgttag gaatctctcc actatcttga aagctattca tttccgagaa catgccacgt 300

gtttcgcaac taaaaatggt atcaaagtaa cagtggaaaa tgcaaagtgt gtgcaagcaa 360

atgcttttat tcaggctgga atatttcagg agtttaaagt tcaggaagag tctgttactt 420

ttcgaattaa tttaactgtc cttttagact gtttatctat ttttggatca agtcctatgc 480

cagggacttt aactgcactt cgaatgtgtt accaaggtta tggttaccct ttgatgctgt 540

tcctggaaga aggaggagtg gtgacagtct gcaaaatcaa tacacaggaa cctgaggaga 600

ccctggactt tgatttctgc agcaccaatg ttattaataa aattattctg cagtcagagg 660

ggctccgtga agcattttct gaattggata tgacgagtga agtcctacaa attaccatgt 720

ctcctgacaa gccttatttc aggttatcta cttttggaaa tgcaggaagt tcccaccttg 780

actatcccaa agattctgat ttgatggaag catttcattg taatcagacc caagtcaaca 840

gatacaagat ttccttactg aaaccctcta caaaggcatt agtcctatct tgtaaggtat 900

ctattcggac agataacaga ggcttccttt cattacagta tatgattaga aatgaagatg 960

gacaaatatg ttttgtggaa tattactgct gccctgatga agaagttcct gaatctgagt 1020

cttgagtatg acaattcact gatatttatg tgtacattta tgatagatga agttcttatt 1080

ctgagtacag tactctttgt catttcatat tggattttct atagagaaga agcacaatgg 1140

ggaagatagg agcaaggtca tgtaccctaa tagttactat gttttgtaaa tccattttgt 1200

agagggcatg taaataaatg ttttcctgta gtcatagatt 1240

<210> 17
<211> 1196
<212> DNA
<213> Mus musculus

<400> 17
gtggaccatc catgcctctc ctaacccagt acaatgaaga ggagtacgaa cagtactgct 60

tagtggccag ccttgacaac gttaggaatc tcttcactgt cttgaaagcc attcatttca 120

gagaacacgc cacgtgtttt gctaccaaaa acggaatcaa ggttacagtg gagaatgcaa 180

agtgtgtgca agcaaatgcc tttattcagg ctgacgtgtt tcaggaattt gtcattcagg 240

aagaatctgt tacttttcga attaacttaa ctatcctttt agactgttta tctatttttg 300

gatcaagtcc tacaccaggg actttgactg cgcttcggat gtgttaccaa ggttatggtc 360

acccactgat gctatttcta gaagaaggag gagtggtgac ggtctgcaaa attaccactc 420

aggagcctga ggagacactg gattttgatt tctgcagcac caatgttatg aataaaatta 480

tcctgcagtc agaggggctc cgggaagcct tttctgagct ggacatgaca ggtgatgtcc 540

tacagatcac tgtgtctcct gacaagccct atttcaggtt gtctactttt ggaaatgcag 600

gaaactccca tcttgactat cccaaagatt ccgacttggt ggaagccttt cactgtgata 660

agacccaggt caacagatac aagctgtcgc tactgaagcc ctctacaaag gcactagctt 720

tatcctgtaa agtgtctatc cggacagata accgaggctt cctctcctta cagtacatga 780

ttagaaatga agatgggcag atatgttttg tggaatatta ctgctgccct gatgaagaag 840

ttcctgagtc ttgaataatt cactgactat ttctgtcttt ttttttttc tttcacgaga 900

cagggtttct ctgtagccct ggctgtcctg gaactcactt tgtagaccac gctggcctcg 960

aactcagaaa ttcacctgcc tctgcctccc gagtgctggg attaaaggca tgcgccacca 1020

cgaccggcga atatttgtgt ctttatagtt ggcgttctca ctctaactct ctaaccttga 1080

tgatttcatt ttggattgta tagagagatg catggagaaa ataggagcaa ggtcagtccc 1140

attctgtaaa tctgtaaata gtttttattt tgtcaataaa tttttcatat agccat 1196

<210> 18
<211> 39
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 18
gcggaattcg gatccaccat gcccccttctg acccaacag 39

<210> 19
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 19
ctcctatctt ccccattgtg 20

<210> 20
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 20
ggtaatttgt aggacttcac 20

<210> 21
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 21
atgttctcgg aaatgaatag 20

<210> 22
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 22
ggggccgaat gcgcagtgga 20

<210> 23
<211> 20
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 23
ctatcccaaa gattctgatt 20

<210> 24
<211> 22
<212> DNA
<213> Artificial Sequence

<220>

<223> Description of Artificial Sequence: primer

<400> 24
aacttcatct atcataaatg ta 22

<210> 25
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 25
aatcagaccc aagtcaacag a 21

<210> 26
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 26
cattggtgct gcagaaatca aa 22

<210> 27
<211> 21
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 27
atgtgttacc aaggttatgg t 21

<210> 28
<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 28
cttagtggcc agccttgaca ac 22

<210> 29

<211> 22
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 29
gcgtgttctc tgaaatgaat gg 22

<210> 30
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: linker
oligonucleotide

<400> 30
ggccgccacc accactttc 19

<210> 31
<211> 19
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: linker
oligonucleotide

<400> 31
tcgagaaagt ggtggtggc 19

<210> 32
<211> 8
<212> PRT
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: gene cassette

<400> 32
Asp Tyr Lys Asp Asp Asp Asp Lys
1 5

<210> 33
<211> 44
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 33
ggatccgaat tcgaatgaag tttcgggcca agatcgtgga cggg 44

<210> 34
<211> 46
<212> DNA
<213> Artificial Sequence

<220>
<223> Description of Artificial Sequence: primer

<400> 34
ggctgcaggt cgactaggac agtgcaggga tgaaatactg aaggga 46

<210> 35
<211> 582
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (3)..(569)

<400> 35
ca gaa acg gaa gca gta agt tcc agc cag gaa atc ccc aca atg cct 47
Glu Thr Glu Ala Val Ser Ser Ser Gln Glu Ile Pro Thr Met Pro
1 5 10 15

cag cca att gag agg gtg tca gtg tca acc cag acc aag aag tta agt 95
Gln Pro Ile Glu Arg Val Ser Val Ser Thr Gln Thr Lys Lys Leu Ser
20 25 30

gcc tct tcc cca cga atg ctg cat cga agc act cag acg acg agc gac 143
Ala Ser Ser Pro Arg Met Leu His Arg Ser Thr Gln Thr Thr Ser Asp
35 40 45

ggc gtc tgt cag agc atg tgc cac gac aag tac acc aaa att ttt aat 191
Gly Val Cys Gln Ser Met Cys His Asp Lys Tyr Thr Lys Ile Phe Asn
50 55 60

gac ttt aaa gac agg atg aag tcg gac cac aag cgt gaa aca gag aga 239
Asp Phe Lys Asp Arg Met Lys Ser Asp His Lys Arg Glu Thr Glu Arg
65 70 75

gtg gtc cgt gaa gcg ctg gaa aag ttg cgt tct gaa atg gaa gaa gaa 287
Val Val Arg Glu Ala Leu Glu Lys Leu Arg Ser Glu Met Glu Glu Glu
80 85 90 95

```
aaa aga caa gcc gta aac aaa gct gtg gcc aac ctg cag ggc gag atg   335
Lys Arg Gln Ala Val Asn Lys Ala Val Ala Asn Leu Gln Gly Glu Met
                100                 105                 110

gac aga aag tgt aag cag gtg aag gaa aag tgc aag gaa gaa ttt gta   383
Asp Arg Lys Cys Lys Gln Val Lys Glu Lys Cys Lys Glu Glu Phe Val
            115                 120                 125

gag gag atc aag aag cta gcg gcg cag cac aag cag ctc att tct cag   431
Glu Glu Ile Lys Lys Leu Ala Ala Gln His Lys Gln Leu Ile Ser Gln
        130                 135                 140

acc aag aag aag cag tgg tgc tac aac tgt gag gag gag gcc atg tac   479
Thr Lys Lys Lys Gln Trp Cys Tyr Asn Cys Glu Glu Glu Ala Met Tyr
    145                 150                 155

cac tgc tgc tgg aac acg tcc tac tgc tcc atc aag tgc cag caa gag   527
His Cys Cys Trp Asn Thr Ser Tyr Cys Ser Ile Lys Cys Gln Gln Glu
160                 165                 170                 175

cac tgg cac gca gag cac aag cgc acc tgc cgc cgg aag aga           569
His Trp His Ala Glu His Lys Arg Thr Cys Arg Arg Lys Arg
                180                 185

tgagcctgcc gcc                                                   582
```

<210> 36
<211> 189
<212> PRT
<213> Mus musculus

<400> 36

```
Glu Thr Glu Ala Val Ser Ser Ser Gln Glu Ile Pro Thr Met Pro Gln
  1               5                  10                  15

Pro Ile Glu Arg Val Ser Val Ser Thr Gln Thr Lys Lys Leu Ser Ala
             20                  25                  30

Ser Ser Pro Arg Met Leu His Arg Ser Thr Gln Thr Thr Ser Asp Gly
         35                  40                  45

Val Cys Gln Ser Met Cys His Asp Lys Tyr Thr Lys Ile Phe Asn Asp
     50                  55                  60

Phe Lys Asp Arg Met Lys Ser Asp His Lys Arg Glu Thr Glu Arg Val
 65                  70                  75                  80

Val Arg Glu Ala Leu Glu Lys Leu Arg Ser Glu Met Glu Glu Glu Lys
                 85                  90                  95

Arg Gln Ala Val Asn Lys Ala Val Ala Asn Leu Gln Gly Glu Met Asp
            100                 105                 110
```

```
Arg Lys Cys Lys Gln Val Lys Glu Lys Cys Lys Glu Glu Phe Val Glu
        115                 120                 125

Glu Ile Lys Lys Leu Ala Ala Gln His Lys Gln Leu Ile Ser Gln Thr
    130                 135                 140

Lys Lys Lys Gln Trp Cys Tyr Asn Cys Glu Glu Glu Ala Met Tyr His
145                 150                 155                 160

Cys Cys Trp Asn Thr Ser Tyr Cys Ser Ile Lys Cys Gln Gln Glu His
                165                 170                 175

Trp His Ala Glu His Lys Arg Thr Cys Arg Arg Lys Arg
            180                 185
```

<210> 37
<211> 433
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (3)..(431)

<400> 37

```
gc atc ccg cag gcc att ctg ggg atg gat gtc ctg tgc cag gcc aag    47
   Ile Pro Gln Ala Ile Leu Gly Met Asp Val Leu Cys Gln Ala Lys
     1               5                  10                  15

tca ggc atg gga aaa aca gca gtg ttt gtc ctg gcc aca ctg cag cag   95
Ser Gly Met Gly Lys Thr Ala Val Phe Val Leu Ala Thr Leu Gln Gln
                 20                  25                  30

ctg gag ccc gtt act ggg cag gtg tct gtg ctg gtg atg tgt cac act   143
Leu Glu Pro Val Thr Gly Gln Val Ser Val Leu Val Met Cys His Thr
             35                  40                  45

agg gag ctg gct ttt cag atc agc aag gaa tat gag cgc ttc tct aag   191
Arg Glu Leu Ala Phe Gln Ile Ser Lys Glu Tyr Glu Arg Phe Ser Lys
         50                  55                  60

tac atg ccg aat gtc aag gtg gca gtg ttt ttt ggc ggt ctg tct atc   239
Tyr Met Pro Asn Val Lys Val Ala Val Phe Phe Gly Gly Leu Ser Ile
     65                  70                  75

aag aag gac gaa gag gtg ctg aag aag aac tgt cca cac atc gtc gtg   287
Lys Lys Asp Glu Glu Val Leu Lys Lys Asn Cys Pro His Ile Val Val
 80                  85                  90                  95

ggg act cct ggc cga att cta gcc ctg gct cga aat aag agc ctg aac   335
Gly Thr Pro Gly Arg Ile Leu Ala Leu Ala Arg Asn Lys Ser Leu Asn
                100                 105                 110
```

```
ctc aaa cac att aaa cac ttt att ttg gac gag tgt gac aag atg ctt   383
Leu Lys His Ile Lys His Phe Ile Leu Asp Glu Cys Asp Lys Met Leu
            115                 120                 125

gaa cag ctc gac atg cgt cgg gat gtc cag gaa att ttt cgc atg acc   431
Glu Gln Leu Asp Met Arg Arg Asp Val Gln Glu Ile Phe Arg Met Thr
        130                 135                 140

ca                                                                433
```

<210> 38
<211> 143
<212> PRT
<213> Mus musculus

<400> 38

```
Ile Pro Gln Ala Ile Leu Gly Met Asp Val Leu Cys Gln Ala Lys Ser
  1               5                  10                  15

Gly Met Gly Lys Thr Ala Val Phe Val Leu Ala Thr Leu Gln Gln Leu
             20                  25                  30

Glu Pro Val Thr Gly Gln Val Ser Val Leu Val Met Cys His Thr Arg
         35                  40                  45

Glu Leu Ala Phe Gln Ile Ser Lys Glu Tyr Glu Arg Phe Ser Lys Tyr
     50                  55                  60

Met Pro Asn Val Lys Val Ala Val Phe Phe Gly Gly Leu Ser Ile Lys
 65                  70                  75                  80

Lys Asp Glu Glu Val Leu Lys Lys Asn Cys Pro His Ile Val Val Gly
                 85                  90                  95

Thr Pro Gly Arg Ile Leu Ala Leu Ala Arg Asn Lys Ser Leu Asn Leu
            100                 105                 110

Lys His Ile Lys His Phe Ile Leu Asp Glu Cys Asp Lys Met Leu Glu
        115                 120                 125

Gln Leu Asp Met Arg Arg Asp Val Gln Glu Ile Phe Arg Met Thr
    130                 135                 140
```

<210> 39
<211> 556
<212> DNA
<213> Mus musculus

<220>
<221> CDS
<222> (2)..(433)

<400> 39

```
g caa gag gct gcg act gta ggg gtg tgt tgg tgc agg tca ggg tta aga 49
  Gln Glu Ala Ala Thr Val Gly Val Cys Trp Cys Arg Ser Gly Leu Arg
    1               5                  10                  15

gac gca ggc atg ctg cgt ccc aag gct ttg acg cag gtg cta agc caa   97
Asp Ala Gly Met Leu Arg Pro Lys Ala Leu Thr Gln Val Leu Ser Gln
             20                  25                  30

gcc aac act gga gga gtc cag agc acc ctg ctg ctg aat aat gag gga   145
Ala Asn Thr Gly Gly Val Gln Ser Thr Leu Leu Leu Asn Asn Glu Gly
         35                  40                  45

tcg ctg ctg gcc tac tcc ggt tat ggg gac aca gat gcc cgg gtc act   193
Ser Leu Leu Ala Tyr Ser Gly Tyr Gly Asp Thr Asp Ala Arg Val Thr
     50                  55                  60

gcg gcc atc gcc agt aac atc tgg gcc gcg tat gat agg aac ggg aac   241
Ala Ala Ile Ala Ser Asn Ile Trp Ala Ala Tyr Asp Arg Asn Gly Asn
 65                  70                  75                  80

caa gcg ttt aat gaa gac agt ctc aaa ttt atc ctg atg gac tgc atg   289
Gln Ala Phe Asn Glu Asp Ser Leu Lys Phe Ile Leu Met Asp Cys Met
                 85                  90                  95

gag ggc cgt gta gcc att acg agg gtg gcc aac ctt ctg cta tgt atg   337
Glu Gly Arg Val Ala Ile Thr Arg Val Ala Asn Leu Leu Leu Cys Met
            100                 105                 110

tat gcc aag gag acc gta ggc ttc gga atg ctc aag gct aag gcc cag   385
Tyr Ala Lys Glu Thr Val Gly Phe Gly Met Leu Lys Ala Lys Ala Gln
        115                 120                 125

gcc ctg gtg cag tac ctg gag gaa ccc ctc acc caa gta gca gca tca   433
Ala Leu Val Gln Tyr Leu Glu Glu Pro Leu Thr Gln Val Ala Ala Ser
    130                 135                 140

taacagcgtg atggaggctg gagtagaaaa gggatgatga tctggaggga ggggcggggc 493

catagaaacg ccatatcggg cgaggtacag gaaggggggg ttgctttttt ctgaataaat 553

ttt                                                               556
```

<210> 40
<211> 144
<212> PRT
<213> Mus musculus

<400> 40

```
Gln Glu Ala Ala Thr Val Gly Val Cys Trp Cys Arg Ser Gly Leu Arg
  1               5                  10                  15

Asp Ala Gly Met Leu Arg Pro Lys Ala Leu Thr Gln Val Leu Ser Gln
             20                  25                  30
```

```
Ala Asn Thr Gly Gly Val Gln Ser Thr Leu Leu Leu Asn Asn Glu Gly
          35                  40                  45

Ser Leu Leu Ala Tyr Ser Gly Tyr Gly Asp Thr Asp Ala Arg Val Thr
      50                  55                  60

Ala Ala Ile Ala Ser Asn Ile Trp Ala Ala Tyr Asp Arg Asn Gly Asn
 65                  70                  75                  80

Gln Ala Phe Asn Glu Asp Ser Leu Lys Phe Ile Leu Met Asp Cys Met
                 85                  90                  95

Glu Gly Arg Val Ala Ile Thr Arg Val Ala Asn Leu Leu Leu Cys Met
            100                 105                 110

Tyr Ala Lys Glu Thr Val Gly Phe Gly Met Leu Lys Ala Lys Ala Gln
        115                 120                 125

Ala Leu Val Gln Tyr Leu Glu Glu Pro Leu Thr Gln Val Ala Ala Ser
    130                 135                 140
```

<210> 41
<211> 662
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (3)..(503)

<400> 41

```
gg cga aga aac tac aac tcc cag ggc gtc ccg gag cag gcc aac ggg    47
   Arg Arg Asn Tyr Asn Ser Gln Gly Val Pro Glu Gln Ala Asn Gly
     1               5                  10                  15

act acg gga agc agc ggg cag cgg ccc gcg gga ggc acc tcg gag atc   95
Thr Thr Gly Ser Ser Gly Gln Arg Pro Ala Gly Gly Thr Ser Glu Ile
                 20                  25                  30

tgg gtg caa aag ccc agg gtt agg aac cgt agc atg ctg cgc ccc aag   143
Trp Val Gln Lys Pro Arg Val Arg Asn Arg Ser Met Leu Arg Pro Lys
             35                  40                  45

gct ttg acc cag gtg cta agc caa gcc aac act gga ggc gtc cag agc   191
Ala Leu Thr Gln Val Leu Ser Gln Ala Asn Thr Gly Gly Val Gln Ser
         50                  55                  60

acc ctg ctg ctg aat aac gag gga tca ctg ctg gcc tac tct ggt tac   239
Thr Leu Leu Leu Asn Asn Glu Gly Ser Leu Leu Ala Tyr Ser Gly Tyr
     65                  70                  75

ggg gac act gac gcc cgg gtc acc gct gcc ata gcc agt aac atc tgg   287
Gly Asp Thr Asp Ala Arg Val Thr Ala Ala Ile Ala Ser Asn Ile Trp
 80                  85                  90                  95
```

```
gcc gcc tac gac cgg aac ggg aac caa gcg ttt aat gaa gac aat ctc   335
Ala Ala Tyr Asp Arg Asn Gly Asn Gln Ala Phe Asn Glu Asp Asn Leu
                100                 105                 110

aaa ttc atc ctc atg gac tgc atg gag ggc cgt gta gcc atc acc cga   383
Lys Phe Ile Leu Met Asp Cys Met Glu Gly Arg Val Ala Ile Thr Arg
            115                 120                 125

gtg gcc aac ctt ctg ctg tgt atg tat gcc aag gag acc gtg ggc ttt   431
Val Ala Asn Leu Leu Leu Cys Met Tyr Ala Lys Glu Thr Val Gly Phe
        130                 135                 140

gga atg ctc aag gcc aag gcc cag gct ttg gtg cag tac ctg gag gag   479
Gly Met Leu Lys Ala Lys Ala Gln Ala Leu Val Gln Tyr Leu Glu Glu
    145                 150                 155

ccc ctc acc caa gtg gcg gca tct taacggcatt ggtggaagct ggggtcagaa  533
Pro Leu Thr Gln Val Ala Ala Ser
160                 165

aagagaaatg accatttgga gggcggggcc tcctagaaga accttcttag acaatggggg 593

gaggatggga ctttgttttt tccaagaata aacttcaact cctgtcaaaa aaaaaaaaaa 653

aagtcgacc                                                        662
```

<210> 42
<211> 167
<212> PRT
<213> Homo sapiens

<400> 42

```
Arg Arg Asn Tyr Asn Ser Gln Gly Val Pro Glu Gln Ala Asn Gly Thr
  1               5                  10                  15

Thr Gly Ser Ser Gly Gln Arg Pro Ala Gly Gly Thr Ser Glu Ile Trp
             20                  25                  30

Val Gln Lys Pro Arg Val Arg Asn Arg Ser Met Leu Arg Pro Lys Ala
         35                  40                  45

Leu Thr Gln Val Leu Ser Gln Ala Asn Thr Gly Gly Val Gln Ser Thr
     50                  55                  60

Leu Leu Leu Asn Asn Glu Gly Ser Leu Leu Ala Tyr Ser Gly Tyr Gly
 65                  70                  75                  80

Asp Thr Asp Ala Arg Val Thr Ala Ala Ile Ala Ser Asn Ile Trp Ala
                 85                  90                  95

Ala Tyr Asp Arg Asn Gly Asn Gln Ala Phe Asn Glu Asp Asn Leu Lys
            100                 105                 110
```

```
Phe Ile Leu Met Asp Cys Met Glu Gly Arg Val Ala Ile Thr Arg Val
        115                 120                 125

Ala Asn Leu Leu Leu Cys Met Tyr Ala Lys Glu Thr Val Gly Phe Gly
    130                 135                 140

Met Leu Lys Ala Lys Ala Gln Ala Leu Val Gln Tyr Leu Glu Glu Pro
145                 150                 155                 160

Leu Thr Gln Val Ala Ala Ser
                165
```

<210> 43
<211> 429
<212> DNA
<213> Homo sapiens

<220>
<221> CDS
<222> (3)..(428)

<400> 43

```
gg gag caa gag gct gcg act gta ggg gtg tgt tgg tgc agg tca ggg    47
   Glu Gln Glu Ala Ala Thr Val Gly Val Cys Trp Cys Arg Ser Gly
     1               5                  10                  15

tta aga gac gca ggc atg ctg cgt ccc aag gct ttg acg cag gtg cta   95
Leu Arg Asp Ala Gly Met Leu Arg Pro Lys Ala Leu Thr Gln Val Leu
                20                  25                  30

agc caa gcc aac act gga gga gtc cag agc acc ctg ctg ctg aat aat   143
Ser Gln Ala Asn Thr Gly Gly Val Gln Ser Thr Leu Leu Leu Asn Asn
            35                  40                  45

gag gga tcg ctg ctg gcc tac tcc ggt tat ggg gac aca gat gcc cgg   191
Glu Gly Ser Leu Leu Ala Tyr Ser Gly Tyr Gly Asp Thr Asp Ala Arg
        50                  55                  60

gtc act gcg gcc atc gcc agt aac atc tgg gcc gcg tat gat agg aac   239
Val Thr Ala Ala Ile Ala Ser Asn Ile Trp Ala Ala Tyr Asp Arg Asn
    65                  70                  75

ggg aac caa gcg ttt aat gaa gac agt ctc aaa ttt atc ctg atg gac   287
Gly Asn Gln Ala Phe Asn Glu Asp Ser Leu Lys Phe Ile Leu Met Asp
80                  85                  90                  95

tgc atg gag ggc cgt gta gcc att acg agg gtg gcc aac ctt ctg cta   335
Cys Met Glu Gly Arg Val Ala Ile Thr Arg Val Ala Asn Leu Leu Leu
                100                 105                 110

tgt atg tat gcc aag gag acc gta ggc ttc gga atg ctc aag gct aag   383
Cys Met Tyr Ala Lys Glu Thr Val Gly Phe Gly Met Leu Lys Ala Lys
            115                 120                 125
```

```
gcc cag gcc ctg gtg cag tac ctg gag gaa ccc ctc acc caa gta g      429
Ala Gln Ala Leu Val Gln Tyr Leu Glu Glu Pro Leu Thr Gln Val
        130                 135                 140
```

<210> 44
<211> 142
<212> PRT
<213> Homo sapiens

<400> 44

```
Glu Gln Glu Ala Ala Thr Val Gly Val Cys Trp Cys Arg Ser Gly Leu
  1               5                  10                  15

Arg Asp Ala Gly Met Leu Arg Pro Lys Ala Leu Thr Gln Val Leu Ser
             20                  25                  30

Gln Ala Asn Thr Gly Gly Val Gln Ser Thr Leu Leu Leu Asn Asn Glu
         35                  40                  45

Gly Ser Leu Leu Ala Tyr Ser Gly Tyr Gly Asp Thr Asp Ala Arg Val
     50                  55                  60

Thr Ala Ala Ile Ala Ser Asn Ile Trp Ala Ala Tyr Asp Arg Asn Gly
 65                  70                  75                  80

Asn Gln Ala Phe Asn Glu Asp Ser Leu Lys Phe Ile Leu Met Asp Cys
                 85                  90                  95

Met Glu Gly Arg Val Ala Ile Thr Arg Val Ala Asn Leu Leu Leu Cys
            100                 105                 110

Met Tyr Ala Lys Glu Thr Val Gly Phe Gly Met Leu Lys Ala Lys Ala
        115                 120                 125

Gln Ala Leu Val Gln Tyr Leu Glu Glu Pro Leu Thr Gln Val
    130                 135                 140
```

Applicant's or agent's file reference	27866/35406	International application No.	PCT/US99/06714

INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13*bis*)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 8 and 20 lines 24 and 31-32

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution

American Type Culture Collection

Address of depositary institution *(including postal code and country)*

10801 University Boulevard

Manassas, VA 20110

Date of deposit	Accession Number
26 March 1999	HB-12689 and HB-12690

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only

☐ This sheet was received with the international application

Authorized officer

For International Bureau use only

☐ This sheet was received by the International Bureau on:

27 MAY 1999

Authorized officer

A. Nickol

Form PCT/RO/134 (July 1998)

INTERNATIONAL SEARCH REPORT

	International Application No PCT/US 99/06714

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/12 C07K14/47 C07K16/18 G01N33/50

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	HILLIER L. ET AL.: "The WashU-Merck EST Project, zk10c07.s1 Soares pregnant uterus NbHPU Homo sapiens cDNA clone 470124 3' similar to SW:RAD1_SCHPO P22193 DNA REPAIR PROTEIN RAD1" EMBL DATABASE ENTRY HSA29300; ACCESSION NUMBER AA029300,25 June 1997 (1997-06-25), XP002111135 cited in the application please compare bp 846-576 of SEQ ID NO:1 with bp 214-484 of AA029300 --- -/--	1-13

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 August 1999	19/08/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Schönwasser, D

6

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/06714

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
X	HILLIER L ET AL: "The WashU-NCI human EST Project, ab35b05.r1 Homo sapiens cDNA clone 842769 5' similar to SW:RAD1_SCHPO P22193 DNA REPAIR PROTEIN RAD1" MATERIALS SCIENCE AND ENGINEERING B, XP002111462 ISSN: 0921-5107 please compare bp 412-842 of SEQ ID NO:1 with bp 1-430 of AA486301 ---	1-13
X	HILLIER L. ET AL.: "The WashU-Merck EST Project 1997, zr56h11.r1 NhHMPu S1 Homo sapiens cDNA clone 667461 5'" EMBL DATABASE ENTRY HS1147325; ACCESSION NUMBER AA227739, 12 February 1997 (1997-02-12), XP002111463 cited in the application please compare bp 1-250 of SEQ ID NO:1 with bp 176-423 of AA227739 ---	5,10
X	MARRA M. ET AL.: "The WashU-HHMI Mouse EST Project, mx23e05.r1 Soares mouse NML Mus musculus cDNA clone 681056 5' " EMBL DATABASE ENTRY MM1165167, ACCESSION NUMBER AA250088 , 15 March 1997 (1997-03-15), XP002111492 please compare bp 29-429 of SEQ ID NO:43 with bp 1-401 of AA250088 ---	36-42,44
X	MARRA M. ET AL.: "The WashU-HHMI Mouse EST Project, va52g09.r1 Soares mouse 3NME12 5 Mus musculus cDNA clone 735040 5' " EMBL DATABASE ENTRY MMAA60182, ACCESSION NUMBER AA260182, - 19 March 1997 (1997-03-19) XP002111493 please compare bp 21-429 of SEQ ID NO:43 with bp 1-405 of AA260182 ---	36-42,44
E	WO 99 31234 A (FREIRE RAIMUNDO ;CANCER RES CAMPAIGN TECH (GB); JACKSON STEPHEN PH) 24 June 1999 (1999-06-24) claims 1-6,8,10-13,15-25 ---	1-8, 10-16, 19-24,30
P,X	FREIRE R ET AL: "Human and mouse homologs of Schizosaccharomyces pombe rad1+ and Saccharomyces cerevisiae RAD17: linkage to checkpoint control and mammalian meiosis" GENES AND DEVELOPMENT, vol. 12, no. 12, 15 August 1998 (1998-08-15), pages 2560-2573 , XP002098400 ISSN: 0890-9369 figures 1,5 --- -/--	1-8, 10-14

6

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 99/06714

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	PARKER A. E. ET AL.: "A Human Homologue of the Schizosaccharomyces pombe rad1+ Checkpoint Gene Encodes an Exonuclease" THE JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 29, no. 17, 17 July 1998 (1998-07-17), pages 18332-18339, XP002111392 page 18334, column 2, line 28 - line 67; figure 1 ---	1-8, 10-13
P,X	UDELL C ET AL: "HRAD1 and MRAD1 encode mammalian homologues of the fission yeast rad1+ cell cycle checkpoint control gene" NUCLEIC ACIDS RESEARCH, vol. 26, no. 17, 1 September 1998 (1998-09-01), pages 3971-3976, XP002098401 ISSN: 0305-1048 page 3973, column 1, line 17 - line 37; figure 1 -----	1-8, 10-13

6

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 99/06714

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. ☐ Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

2. ☒ Claims Nos.: 23,28,29,34,35
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:
see FURTHER INFORMATION sheet PCT/ISA/210

3. ☐ Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. ☐ As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. ☐ As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. ☐ As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. ☐ No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

☐ The additional search fees were accompanied by the applicant's protest.

☐ No protest accompanied the payment of additional search fees.

FURTHER INFORMATION CONTINUED FROM PCT/ISA/ 210

Continuation of Box I.2

Claims Nos.: 23 28 29 34 35

It is not possible to carry out a meaningful search for claims 23, 28, 29, 34 and 35, since the claimed subject-matter is insufficiently described.

The applicant's attention is drawn to the fact that claims relating to inventions in respect of which no international search report has been established need not be the subject of an international preliminary examination (Rule 66.1(e) PCT). The applicant is advised that the EPO policy when acting as an International Preliminary Examining Authority is normally not to carry out a preliminary examination on matter which has not been searched. This is the case irrespective of whether or not the claims are amended following receipt of the search report or during any Chapter II procedure.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/US 99/06714

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9931234 A	24-06-1999	GB 2333776 A	04-08-1999